

01170
1



**UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO**

**FACULTAD DE INGENIERÍA
DIVISIÓN DE ESTUDIOS DE
POSGRADO**

**SELECCIÓN DE WAVELET
EN EL ESTÁNDAR
JPEG 2000**

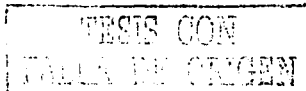
T E S I S

Que para obtener el grado de:
**MAESTRÍA EN INGENIERÍA
ELÉCTRICA**
con especialidad en
PROCESAMIENTO DIGITAL DE SEÑALES

presenta
**MARÍA PATRICIA CAMARILLO
SANDOVAL**

Director de tesis:
Dr. Boris Escalante Ramírez

México, D.F.



2003



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

*A Mamámay,
tu recuerdo me acompañará
por siempre ...*

Autorizo a la Dirección General de Bibliotecas
UNAM a difundir en formato electrónico e impreso
el contenido de mi trabajo receptor

NOMBRE: María Fabrice
Camarillo Sandoval

FECHA: 27 febrero - 2003

FIRMA: M. Fabrice S.

TESIS CON
FALLA DE ORIGEN

0

Agradecimientos.

A mis padres Paty y Manolo con los que siempre puedo contar.

A mis hermanos, Manolo, Pily y Ale; por su cariño, ayuda y comprensión en todo momento y por ser lo mejor de mi vida.

A mi tía Chabe que siempre ha demostrado que cree en mí.

A mi primo Pepe, por su cariño incondicional.

Al Dr. Boris Escalante por su apoyo y asesoría, pero principalmente porque nunca permitió que desistiera.

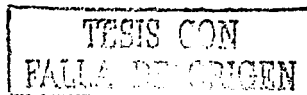
A Alma, Diana, Mauro, Paco y Soco por su apoyo y amistad.

A los profesores que compartieron su invaluable conocimiento conmigo, principalmente al Dr. Francisco García Ugalde y al Dr. Miguel Moctezuma Flores.

A CONACYT porque sin su ayuda no hubiera alcanzado esta meta.

Es bueno agradecer a las personas a las que les importamos, porque precisamente el apoyo y comprensión de nuestra familia y amigos son lo que nos impulsan a llevar a cabo nuestros sueños. Por todos ellos agradezco a Dios y por permitirme culminar esta etapa de mi vida.

María Patricia Camarillo Sandoval.



ÍNDICE

Capítulo 1. Introducción

1.1 Aplicación de la compresión de imágenes.....	1
1.2 Definición del problema	2
1.3 Justificación e Hipótesis.....	4
1.4 Objetivos de la tesis	6
1.5 Aportaciones	6
1.6 Descripción de los capítulos.....	7

Capítulo 2. Antecedentes

2.1 Transformación wavelet.....	9
2.1.1 Definición de wavelet	9
2.1.2 Transformaciones STFT y CWT.....	10
2.1.3 Codificación sub-banda.....	11
2.1.4 Discretización de la transformada wavelet	12
2.1.5 Análisis Multi-resolución.....	15
2.1.6 Esquema Lifting.....	16
2.1.7 Familias de wavelets	18
2.1.7.1 Wavelet Haar	18
2.1.7.2 Wavelet Shannon.....	18
2.1.7.3 Wavelets de soporte infinito	18
2.1.7.4 Wavelets ortonormales de soporte finito.....	20
2.1.7.5 Wavelet biortogonales	21
2.1.8 Estructuras de descomposición	23
2.2 Cuantización.....	25
2.2.1 Escalar	25
2.2.2 TCQ.....	26

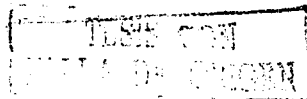
2.3 Codificación por entropía.....	27
2.3.1 Codificador MQ.....	28
2.3.2 EZW.....	29
2.3.3 Codificación embebida de bloques.....	31

Capítulo 3. JPEG 2000

3.1 Objetivo del estándar.....	35
3.2 Partes del JPEG 2000.....	35
3.3 Características generales.....	36
3.4 Esquema general del JPEG 2000.....	37
3.5 Arquitectura.....	38
3.5.1 Pre-procesamiento.....	38
3.5.2 Procesamiento principal.....	40
3.6 Parte II del estándar.....	46
3.7 Notas sobre el estándar.....	47
3.8 Glosario.....	47

Capítulo 4. Selección de wavelets.

4.1 Objetivo de la selección de wavelets.....	48
4.2 Metodología de evaluación.....	48
4.3 Selección del grupo de estudio.....	49
4.3.1 Variables consideradas.....	49
4.3.2 Efecto de las variables en el desempeño de la wavelet.....	49
4.3.3 Análisis de las wavelets comparadas.....	54



Capítulo 5. Resultados

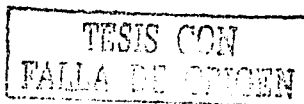
5.1 Software utilizado	64
5.2 Variables que se fijaron en los experimentos	64
5.3 Resultados	66
5.3.1 Análisis objetivo	67
5.3.2 Análisis subjetivo	67
5.3.3 Comparación entre la calidad subjetiva y objetiva	68

Capítulo 6. Conclusiones

6.1 Discusión de los resultados cuantitativos	86
6.2 Discusión de los resultados cualitativos	87
6.3 Conclusiones	90
6.4 Investigación futura	94

Referencias96

Resumen99



Capítulo 1

INTRODUCCIÓN

1.1 Aplicación de la compresión de imágenes.

En nuestros días es frecuente la escena de un sistema de cómputo doméstico al que se encuentra conectado un escáner e inclusive una cámara digital, si agregamos a este panorama que el sistema anteriormente descrito no se encuentra aislado, sino en conexión con el resto del mundo a través del Internet; la importancia de algoritmos eficientes para la compresión de datos multimedia, antes de su almacenamiento y transmisión, se hace evidente e imprescindible dadas las limitaciones físicas de ancho de banda y costo de almacenaje. Las aplicaciones que han surgido a partir de esta apertura ya no se limitan a fines científicos donde imágenes de satélite o de telescopios espaciales son transmitidas y almacenadas para su análisis, ahora podemos hablar desde diagnósticos médicos a distancia o videoconferencias hasta la simple transmisión de imágenes para publicidad o diversión.

El gusto natural que tenemos por las imágenes nos fuerza a ser cada vez más ambiciosos en cuanto al ahorro de bits se trata, lo que desemboca en la búsqueda de la manera más económica para representar escenas que expresen hechos o pensamientos y que adicionalmente nos permita la fácil manipulación de la información visual. Ante este mercado se han desarrollado una amplia variedad de técnicas para la compresión de imágenes estáticas y de video, las cuales en esencia buscan eliminar la redundancia espacial, espectral y temporal presente en las escenas; expresando la información significativa en función de descriptores óptimos.

Independiente a esta necesidad, pero derivada de la misma fuente, se encuentra el problema de comunicación entre los sistemas digitales, aquí es donde el papel de los estándares es notorio. En el panorama descrito en las primeras líneas de este capítulo, las características del sistema de cómputo no fueron precisas y en la vida real esta información es entregada a las aplicaciones hasta cuando el sistema hace uso de ellas. Para que este sistema no quede excluido se requiere el concepto de interoperabilidad, lo que permite la comunicación entre nuestro sistema y el resto de los conectados a Internet con base en el establecimiento de protocolos, formatos y algoritmos especializados que sean capaces de trabajar en cualquier ambiente: El JPEG* 2000 es la nueva respuesta a este requerimiento.

En este esquema, las dos clases de compresión siguen siendo válidas y completamente dependientes de la aplicación de la que se trate. Pero hablando de la interoperabilidad, el problema de la escasez de recursos se resuelve con el concepto de la transmisión

* Siglas de Joint Photographic Experts Group que en extensión se refieren a los estándares emitidos por este comité.

PAGINACION DISCONTINUA

progresiva, con el cual logramos obtener la compresión sin pérdida (también llamada reversible), contando inicialmente con la información de la compresión con pérdida (también llamada irreversible). Para este efecto como se verá más adelante, el estándar utiliza términos como escalabilidad de calidad y de resolución; mientras que para implantar el proceso reversible se desarrolló un nuevo esquema de filtrado llamado lifting, que será explicado en el siguiente capítulo.

Con las condiciones expuestas anteriormente, lo que se espera de los algoritmos de compresión es:

- Una muy baja tasa de bit para las primeras aproximaciones donde se tiene que tomar la decisión entre retener o rechazar la imagen que se está presentando.
- Obtener el máximo provecho de los datos transmitidos previamente, con la finalidad de minimizar la información adicional requerida en la siguiente etapa.
- Presentar la capacidad de transmitir completamente la imagen con una baja tasa de bit global y con excelente calidad (incluso sin pérdida).
- Finalmente, que los algoritmos de codificación / decodificación sean rápidos y apropiados para las implantaciones en hardware.

1.2 Definición del problema.

El impacto del JPEG 2000 y los cambios que propone marcando una tendencia bien definida para las aplicaciones actuales del Internet, videoconferencias y las telecomunicaciones en general, han hecho que sea considerado incluso en el estándar de compresión de video MPEG[†]-4. El estándar pues se ha establecido, por lo que las herramientas para el almacenamiento y transmisión de imágenes deben apegarse a él, sin embargo, aunque los algoritmos quedan definidos existen parámetros que determinan el performance global del proceso de compresión; este contexto es el que envuelve el problema que me ocupa.

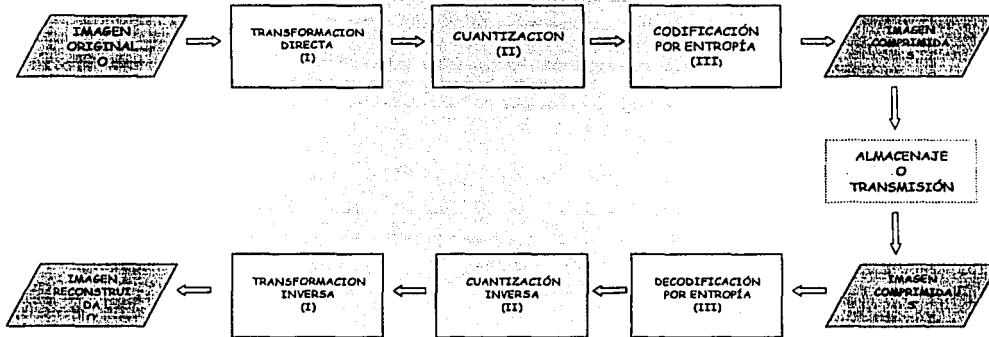
La figura 1.1 presenta el esquema de los codificadores basados en transformada, entre los que se cuenta con el JPEG 2000, el cual puede dividirse en tres componentes principales: (i) Descomposición de la imagen usando transformaciones; (ii) Cuantización de los coeficientes transformados; (iii) Codificación basada en entropía de los coeficientes cuantizados.

Aunque el esquema completo es tratado a lo largo de los siguientes capítulos, la atención se centrará en el primer componente y específicamente en la transformación wavelet. En esta categoría de codificadores la selección apropiada de la wavelet es crucial para la determinación del performance de la compresión. Las tendencias actuales se centran en los procedimientos de búsqueda de la base óptima (tipo, orden y árbol) y el JPEG 2000 no es

[†] Siglas de Moving Picture Experts Group.

rígido en este sentido, permitiendo la libre selección del núcleo de la transformación. La selección de la wavelet es entonces un factor que se puede explotar en el diseño de nuevas aplicaciones.

Fig. 1.1. Esquema general de codificación de imágenes por transformada.



Para hacer un análisis de transformaciones enteros a enteros es necesario que los filtros puedan ser expresados como números racionales donde el denominador sea una potencia de 2, además de que las métricas de desempeño consideran otros criterios, en la cual el PSNR no está incluido. El presente trabajo se limita a la compresión con pérdida, los filtros fueron implementados sin tener cuidado en que las operaciones fueran reversibles, incluso alguno de los filtros que se consideraron tienen coeficientes irracionales.

Entonces el problema que me ocupa puede enunciarse como sigue:

Estudio y determinación de la transformada wavelet óptima para la compresión de imágenes con pérdida, en el contexto del estándar JPEG 2000.

1.3 Justificación e Hipótesis.

Las razones para centrar nuestra atención en el JPEG 2000 se explican por la necesidad de interoperabilidad de los sistemas actuales; independientemente a este estándar los esquemas que han surgido para la compresión de imágenes utilizando wavelets abarcan desde simples codificadores por entropía, hasta técnicas que involucran cuantización vectorial, codificación de árboles jerárquicos e incluso codificación basada en bordes. Una de las características que hace tan popular el uso de las wavelets es que representan una herramienta biológicamente plausible, dado que la base de funciones que utiliza permite la localización de la señal tanto en el dominio del tiempo como en el de la frecuencia y, el análisis multi-resolución corresponde a una de las características del procesamiento de bajo nivel que descompone a la imagen en versiones a diferentes escalas de la misma, encontrándose semejanzas con el sistema visual humano.

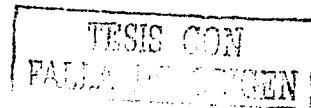
Otras ventajas que presentan las wavelets en su implementación son:

- Muchas funciones wavelets tienen soporte compacto.
- Se pueden crear bases ortogonales, semi-ortogonales o biortogonales con ellas.
- Existen algoritmos rápidos de transformación.
- La transformación se puede implementar a través de bancos de filtros
- Existen fuertes fundamentos matemáticos que permiten la posibilidad de diseñar a la medida la wavelet a través del esquema lifting.

Más allá de las características propias de las wavelets, el JPEG 2000 a través del proceso que sigue se parece más a nuestro sistema visual, en el cual no se almacena la información muestra por muestra[‡] sino que, después de un procesamiento de alto nivel, la información percibida es almacenada en una clase de memoria asociativa contextual y orientada a objetos.

Un campo de estudio sobre transformaciones wavelet se dedica a la búsqueda de propiedades que permitan mejorar la representación de los eventos que ocurren en la imagen. Las características ideales que deben cumplir estas funciones en ocasiones se contraponen entre sí y es cuando se hace necesario decidir cuál transformada se acopla mejor a las necesidades de la aplicación. El gran desarrollo matemático logrado en la última década pone frente a los investigadores y desarrolladores de aplicaciones una amplia gama de funciones, que hasta ahora han sido seleccionadas por los resultados que experimentalmente se han encontrado. Los trabajos enfocados al análisis del desempeño de las transformaciones, o bien se han hecho aisladas del esquema de compresión [A2][C2] o con poca relación al estándar en el que finalmente trabajarán [M3][A1] y que eventualmente será el estándar JPEG 2000.

[‡] Nuestra representación de las imágenes es con pérdida, por naturaleza.



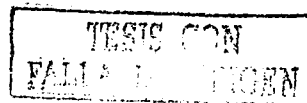
Desde un enfoque en el que la forma de la onda es un factor importante en la búsqueda de la wavelet que cumpla con los requisitos teóricos para tener un desempeño eficiente en la compresión de imágenes, mi interés se centra en las funciones gaussianas, las cuales han demostrado ser eficientes en las aplicaciones como el análisis y eliminación del ruido, ya que al estar íntimamente relacionadas al sistema visual humano son altamente eficientes para representar los eventos de una imagen. Estas razones motivan a pensar en que puedan concentrar en pocos coeficientes la información relevante de la imagen, aunado a esto las funciones gaussianas son altamente regulares incluso pertenecen al espacio de Hölder C^α .

En este campo la transformada Hermitiana y de Gabor son populares. De la primera sabemos que por la teoría de escala, una función puede ser descrita como una superposición de gaussianas, sin embargo, necesitamos un número infinito de tales funciones, las cuales no son linealmente independientes, por lo que no pueden formar una base. Por otro lado, las bases de $L_2(\mathbb{R})$ están formadas de elementos infinitos y contables, así que si tomamos las funciones $\{e^{-n \cdot m^2} : n \geq 0\}$ y las ortonormalizamos usando el método de Gram-Schmidt se obtiene un conjunto contable, discreto de las funciones Hermite (polinomios multiplicando gaussianas), que son suficientes para cubrir $f(t)$. En cuanto a las funciones de Gabor, estas son seleccionadas por su relevancia psicovisual, ya que su forma es parecida a los campos receptivos de algunas células de la corteza visual primaria, son localizables tanto en el dominio del tiempo como en el de la frecuencia, su forma de onda es plana y consiste de senoidales restringidas por una función envolvente gaussiana.

Sin embargo, en el desarrollo de las aplicaciones con wavelets otras características de la función, como los momentos nulos y las características de los filtros asociados, se han considerado más influyentes en el desempeño total del esquema de compresión, permitiendo que la forma de onda sea modificada por estos factores.

Finalmente en esta búsqueda si se desea determinar una transformación "óptima" se debe establecer los criterios que se tomarán en cuenta. Como antecedente, en [M4] se evalúa las transformaciones de funciones 1D empleando el error cuadrático medio y la función tasa-distorsión; en [G4] se evalúa los problemas de implementación de la transformada wavelet utilizando un esquema de árbol jerárquico para su codificación, sin embargo, lo que interesa es el performance global de la compresión por lo que en este caso se utilizará tanto el PSNR como una calificación subjetiva solicitada a sujetos sin conocimientos en procesamiento de señales.

La hipótesis que se sostiene en el presente trabajo es que el emplear funciones wavelet con una forma de onda parecida a la función gaussiana, como principal criterio de selección, mejorará el performance global en la compresión de imágenes tomando como ambiente de comparación el emergente estándar JPEG 2000.



1.4 Objetivos de la tesis.

- Exponer el estado del arte en cuanto a algoritmos de compresión de imágenes por transformada, específicamente por transformada wavelet.
- Describir el estándar para la compresión de imágenes estáticas: JPEG 2000.
- Estudiar las características individuales de los sistemas wavelets más populares y que fueran biológicamente plausibles, tomando como referencia el sistema visual humano.
- Seleccionar un conjunto de sistemas wavelet que cumpliendo con los requerimientos del estándar JPEG 2000, conservarán la forma de onda semejante a la función gaussiana o a sus derivadas, dado que se ha demostrado que son buenos descriptores en el análisis de imágenes.
- Determinar de acuerdo al desempeño de la compresión, las características más ponderables para el diseño de una metodología para la determinación de la transformada wavelet óptima en JPEG 2000.
- Determinar si la wavelet gaussiana es una buena opción para la compresión de imágenes utilizando el estándar JPEG 2000.
- Comprobar teóricamente la mejor selección de la wavelet en el contexto del estándar JPEG 2000.

1.5 Aportaciones.

Para obtener mayor provecho de las bondades de un esquema de compresión debemos comprender cada uno de los procesos involucrados en él, de esta manera podemos mejorarlo ya sea haciendo un correcto uso de la flexibilidad de los parámetros que nos presenta o proponiendo modificaciones a alguno de sus componentes. Este trabajo está enfocado al primer camino.

Sabiendo que, aunque el estándar sugiere transformadas wavelets específicas, no existe imposición para su uso. La tarea fue encontrar el mejor sistema wavelet a partir de su desempeño global en el proceso de compresión. A partir del conjunto de sistemas wavelets empleados en la compresión de imágenes, se seleccionaron aquellos con las forma de onda de las función wavelet más suave y parecida a la función gaussiana, tan relacionada al SVH (Sistema Visual Humano).

En el estudio incluí tanto sistemas ortogonales como biortogonales, identifiqué cada una de las variables que se pueden definir en el diseño del sistema, su impacto en el desempeño del mismo y su relación con la forma de onda de la función wavelet.

La transformación nos permite distinguir la información relevante, la idea es que esta distinción se realice con base a lo que es perceptualmente importante. Haciendo una medición de los resultados obtenidos, se determinó el conjunto de sistemas wavelets que tuvieron mejor desempeño y las características que comparten, las cuales sientan las bases para un mejor diseño de sistemas wavelets dependientes de la aplicación que se trate.

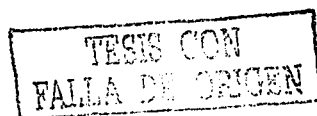
Entre las conclusiones se encuentran las siguientes :

- El desempeño de la compresión depende principalmente de la regularidad de la señal.
- La fase lineal de los filtros es imprescindible para evitar distorsiones en los bordes de la imagen, esta es la razón por la que los sistemas biortogonales son superiores a los ortogonales.
- Dentro de los sistemas biortogonales, los de mejor desempeño fueron aquellos en los que los filtros de análisis se asemejaban a los de síntesis.
- La regularidad es una característica que se prefiere en la síntesis y para el caso de las calificaciones subjetivas la mejor calidad las obtuvieron la wavelet con la forma de onda más regular (los sistemas reversibles) y las wavelets que se basan en funciones B-spline; mientras que por calidad objetiva destacó la segunda. Por otro lado los momentos nulos, asociados con la propiedad de oscilación de la forma de onda es preferible que se presente en el análisis para asegurar la magnitud pequeña de los coeficientes wavelet.
- Las variables más influyentes en las calificaciones objetivas son: (i) Simetría de los filtros y (ii) Número de momentos nulos; mientras que para las calificaciones subjetivas son: (i) Simetría de los filtros, (ii) Regularidad de la función wavelet y (iii) Forma de onda de la wavelet.

1.6 Descripción de los capítulos.

El trabajo está organizado en seis capítulos que presentan la metodología utilizada para llevar a cabo el estudio y la implementación de los experimentos. Este capítulo se centró en el entendimiento del problema, la descripción del contexto en el que se encuentra enmarcado, la presentación de la hipótesis que motivó el desarrollo de este trabajo, los objetivos que determinaron la manera en que éste se realizaría y las aportaciones logradas.

En el capítulo 2 se explicarán los principios teóricos que sustentan cada etapa del proceso de compresión, estableciendo el estado del arte en cuanto a los algoritmos de implantación, así como sus antecedentes más sobresalientes.

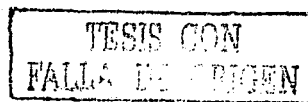


En el capítulo 3 se describirá con algún detenimiento la parte I del estándar JPEG 2000. Se explicarán los algoritmos que se emplean en cada uno de los bloques de la figura 1.1, junto con las libertades y restricciones que imponen en el estándar.

En el capítulo 4 se expondrá la metodología para la selección adecuada de sistemas wavelets dentro del esquema de compresión y se dará a conocer los sistemas que se incluyeron en el estudio.

En el capítulo 5 se describirán los experimentos realizados y se presentará la información de los resultados obtenidos,.

Por último, en el capítulo 6 se discutirán los resultados, se enunciarán las conclusiones obtenidas y se proporcionará un panorama de hacia donde se dirige la investigación en el tema de las transformaciones wavelet.



Capítulo 2

Antecedentes teóricos.

Este capítulo tiene como objetivos:

- Enunciar los fundamentos teóricos en los que se basa este trabajo, para poder ser utilizados en los siguientes capítulos.
- Establecer una nomenclatura que permita el entendimiento de los desarrollos que a continuación se presentan.

2.1 Transformación wavelet

En los métodos tradicionales para la codificación de imágenes, la imagen es dividida en bloques para posteriormente transformarlos y codificarlos independientemente, el ejemplo clásico de esta técnica es el estándar JPEG. Sin embargo, este esquema tiene las peculiaridades de que; (i) no permite obtener provecho de la correlación existente entre los bloques y, (ii) presenta artefactos de bloques que son aún más perceptibles cuando se comprime la imagen a bajas tasas de bits. La alternativa que surgió entonces fue la transformada discreta wavelet que tiene la habilidad de proporcionar una representación multi-resolución de la imagen. Estas representaciones, generalmente piramidales permiten que los errores después de la reconstrucción sean distribuidos sobre toda la imagen, mejorando su calidad visual; la idea fundamental es analizar las imágenes de acuerdo a su escala, a través de una función a la que se le desplaza y cambia de escala con la finalidad de obtener diferentes aproximaciones de la señal original. Como se miden las fluctuaciones promedio de la señal a diferentes escalas, el análisis es menos sensible al ruido.

2.1.1 Definición de wavelet

A principios de los 80's, el geofísico Jean Morlet tuvo la idea de basar el análisis de una señal en una función h , a la que llamó *wavelet*, que fuera bien localizada en tiempo y en frecuencia. Esta función era dilatada y trasladada para formar una familia de funciones de análisis que permitieran un análisis local de las señales, incluso de aquellas que fueran temporales, variantes en el tiempo y no estacionarias.

Desde entonces los campos de investigación donde se emplean wavelets han ido en aumento, haciendo inexistente una definición formal de wavelet. En [S4] Sweldens describe a las wavelets de la siguiente manera:

“Una wavelet es denotada por $\psi_\lambda(x)$ donde x pertenece a una dominio espacial (no definido) X , ψ_λ pertenece a una clase (no definida) F de funciones, y λ pertenece a un conjunto de índices (no definidos) Λ . Además nos referiremos a $\Psi = \{\psi_\lambda | \lambda \in \Lambda\}$ como la base wavelet siempre y cuando cumpla con: (i) Sean bloques de construcción para funciones generales, (ii) Tengan buena localización en el espacio-frecuencia. (iii) Existan algoritmos de transformación rápidos.”

A continuación se introducen los conceptos básicos de la teoría de wavelets.

* En este punto se puede asociar X a la línea real, F a $L_2(\mathfrak{R})$ y Λ a Z^2 ; con $\lambda = (j, l)$ y $\psi_\lambda(x)$ como $2^{j/2} \psi(2^j x - l)$

2.1.2 Transformaciones STFT y CWT

Aunque la teoría de wavelets tiene una cuna francesa con el geofísico J. Morlet y el físico A. Grossmann[†], fueron dos investigadores americanos, Daubechies y Mallat, quienes encontraron la relación de wavelets y el procesamiento digital de señales. Ellos presentaron la construcción de wavelets basada en que la iteración de un banco de filtros en el tiempo discreto converge a una base de wavelet en tiempo continuo.

La transformación wavelet (WT) como una transformación tiempo-frecuencia (aunque es más correcto describirla como una transformación tiempo-escala) es una alternativa directa a la transformada de Fourier de tiempo limitado (STFT) y es una herramienta flexible para la descomposición multi-resolución de funciones continuas. Si se parte de que la transformada de Fourier y su inversa están definidas como:

$$F(\omega) = \int_{-\infty}^{\infty} f(t)(\cos \omega t - j \operatorname{sen} \omega t) dt$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)(\cos \omega t + j \operatorname{sen} \omega t) d\omega$$

donde $F(\omega)$ es la transformada de Fourier de $f(t)$. Se puede ver que $F(\omega)$ es una función que describe la contribución de senos y cosenos a la construcción de la imagen original en el dominio del tiempo. La contribución de estas "funciones base"[‡] está limitada en el tiempo por la duración de la señal a ser analizada, por lo que esta independencia resulta en una descripción puramente en el dominio de la frecuencia.

La transformada de Fourier es ideal para el análisis de señales estacionarias, pero para el caso de las señales no estacionarias se requiere una transformación a un dominio tiempo-frecuencia, para lo cual se emplea la STFT definida por Gabor:

$$STFT(\tau, \omega) = \int s(t)g(t - \tau)e^{-j\omega t} dt$$

la cual representa la transformada de Fourier de una señal $s(t)$, después de habersele aplicado una ventana con la función $g(t)$ alrededor del tiempo τ . Como la función ventana es desplazada en el tiempo sobre toda la señal y, transformaciones consecutivas de porciones de la señal que se traslapan son calculadas, se llega a una descripción del espectro de la señal a través del tiempo, llamado *espectrograma*. Este método supone estacionariedad sobre la ventana, lo que frecuentemente se cumple si la ventana es pequeña.

El tamaño de la ventana afecta directamente la resolución en la frecuencia, ya que el seleccionar una ventana pequeña resulta en un número reducido de muestras para el cálculo de la transformada de Fourier, lo que lleva a un número reducido de frecuencias discretas que pueden ser representadas en el dominio de la frecuencia, en consecuencia, se reduce la posibilidad de discriminar senoidales

[†] En este enfoque se sustituyen la modulación de las exponenciales complejas por operaciones de escalamiento, remplazando la noción de escala por la de frecuencia.

[‡] Este término se refiere a un conjunto completo de funciones que pueden, cuando se combinan en una suma ponderada, usarse para construir una señal dada.

de diferentes frecuencias. A este efecto se le llama principio de incertidumbre y puede ser escrito como:

$$\Delta t \Delta f \geq \frac{1}{4\pi}$$

En el caso de la STFT las funciones base son las exponenciales complejas donde se aplicó una ventana $g(t)$ centrada en τ . Si se denotan a las funciones base, $k_{\tau,\omega}(t)$, se puede expresar la STFT como el siguiente producto interno:

$$STFT(\tau, \omega) = \int s(t) k_{\tau,\omega}(t) dt$$

Tomando la misma notación para la WT, reemplazando la variable de frecuencia por la variable de escala a y la variable de desplazamiento en el tiempo por b :

$$\left. \begin{aligned} CWT(b, a) &= \int s(t) k_{b,a}(t) dt \\ k_{b,a}(t) &= \frac{1}{\sqrt{a}} h\left(\frac{t-b}{b}\right) \end{aligned} \right\} (2.1)$$

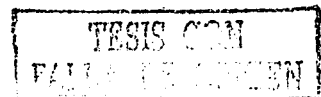
De estas ecuaciones se puede observar que la WT realiza una descomposición de la señal $s(t)$ en una suma ponderada de funciones wavelets de diferentes escalas de $h(t)$. Generalmente $h(t)$ es una función compleja y todas las funciones base parten de esta función común llamada la wavelet madre. El producto punto entre la función a analizar y las funciones $h(t)$ es una representación en tiempo-escala llamada transformada wavelet. Si la wavelet h satisface la condición de admisibilidad (lo que significa que h oscila ya que su integral es nula), entonces la transformada puede ser invertida. Al conjunto de funciones $\{k_{b,a}(t)\}$ se le llama familia wavelet.

Debido al escalamiento, wavelets de altas frecuencias son de duración limitada y las wavelets de bajas frecuencias son relativamente más largas en duración. Esta característica de longitud de ventana variable son obviamente apropiadas para el análisis de funciones cortas de alta frecuencia y componentes duraderos de baja frecuencia.

2.1.3 Codificación sub-banda

El esquema ideal de la codificación sub-banda consiste en filtrar la imagen de entrada en m canales de frecuencia y submuestrear las correspondientes salidas, reteniendo solo un punto en m . Esta operación consiste en restringir una secuencia definida en los enteros al conjunto de los enteros múltiplos de m , a esto se le llama decimación. El esquema para reconstruir la señal original es el dual del esquema de análisis. Se extienden las secuencias de cada canal, insertando 0's en los enteros que no son múltiplos de m , y otra vez se vuelven a procesar las secuencias con los mismos filtros empleados en la etapa de análisis. Sin embargo, este esquema es impráctico ya que requiere que los filtros que se usen para lograr una buena definición de frecuencia sean de longitud infinita. En la práctica, se limita la duración usando alguna función de ponderación que permita que no exista distorsión en la señal reconstruida. La Fig. 2.1 muestra este esquema para $m=2$.

Los filtros QMF fueron estudiados desde 1977 pero no se podía implementar la reconstrucción perfecta hasta que en 1984 el problema de seleccionar $\{h_i\}$ y $\{g_i\}$ de tal manera que se pudiera aproximar a los filtros paso-banda para cubrir aproximadamente todas las frecuencias e igualar la entrada con la salida de los bancos fue resuelta por Esteban y Galand.



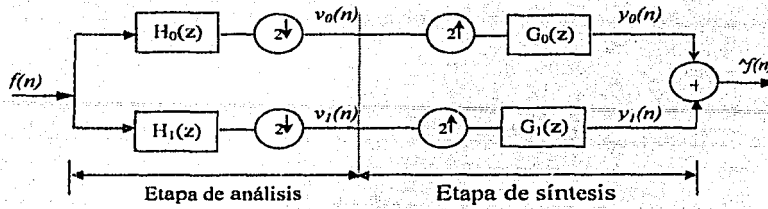


Fig. 2.1 Banco de Filtros de dos bandas.

Las condiciones que deben de cumplir los filtros para que la reconstrucción sea perfecta son [S1, G1]:

$$\left. \begin{aligned} G_0(z) &= -H_1(-z) \\ G_1(z) &= H_0(-z) \\ H_1(z) &= z^{-(N-1)}H_0(-z^{-1}) \\ H_0(z)H_0(z^{-1}) + H_0(-z)H_0(-z^{-1}) &= 1 \end{aligned} \right\} (2.3)$$

Donde N es el número de coeficientes de $H_0(z)$ y $H_1(z)$

Para utilizar filtros de análisis FIR de N coeficientes; $H_0(z)$ y $H_1(z)$ son seguidos por sub-muestreadores de tasa 2. Las señales $v_0(n)$ y $v_1(n)$ son llamadas las sub-bandas paso-bajas y paso-altas, respectivamente. En la etapa de síntesis, las sub-bandas son sobre-muestreadas y entonces interpoladas con los filtros $G_0(z)$ y $G_1(z)$. Finalmente, la señal reconstruida $\hat{f}(n)$ se obtiene al sumar las sub-bandas decimadas e interpoladas.

Las secuencias $\{h_0(n)\}$ y $\{h_1(n)\}$ son las mismas que se definen para la transformación wavelet. Además, Daubechies demostró que un árbol sub-banda de tiempo discreto puede servir como algoritmo rápido de la transformada wavelet si es alimentado con los coeficientes de escalamiento, $\{c_{0,n}\}$ de resolución completa. Esto implica que primero se debe proyectar la función continua $f(t)$ sobre $\{\phi_{0,n}(t)\}$ y posteriormente colocar los coeficientes como entrada al banco de filtros.

$$c_{0,n} = \int_{-\infty}^{+\infty} f(t)\phi(t-n)dt$$

Las características de las funciones ϕ se explicarán en la siguiente sección. Dado que es claro que $f(n)$ es diferente a $c_{0,n}$, no podemos alimentar un banco de filtros sub-banda con una señal discreta y llamarle a esto una transformación wavelet discreta. Hablamos de transformación sub-banda en el caso de que la señal de entrada sólo haya sido muestreada como es el caso de $f(n)$.

2.1.4 Discretización de la transformada wavelet.

La implementación de la DWT puede ser introducida a través de la descomposición sub-banda. La estructura común del banco de filtros puede ser implementada eficientemente con el arreglo de la Figura 2.2. En cada etapa de la estructura, la señal es dividida igualmente en dos componentes de frecuencia, uno bajo y otro alto. Después del filtrado, los dos componentes tienen redundancia y es válido submuestrear los componentes por un factor de 2 sin pérdida de información.

La implementación de esta estructura de árbol de filtros resulta en un banco de filtros contiguos paso-banda cubriendo logarítmicamente todas las frecuencias. Esta estructura de filtros es equivalente a la CWT de una señal muestreada en una serie de valores de escala en octavas. Los wavelets de la WT son equivalentes a las funciones de los filtros paso-banda.

Si la wavelet es repetidamente dilatada por un factor de dos, se logra la cobertura logarítmica de las frecuencias, aunque existen restricciones sobre las wavelets para que exhiban las características QMF en cada nivel de dilatación.

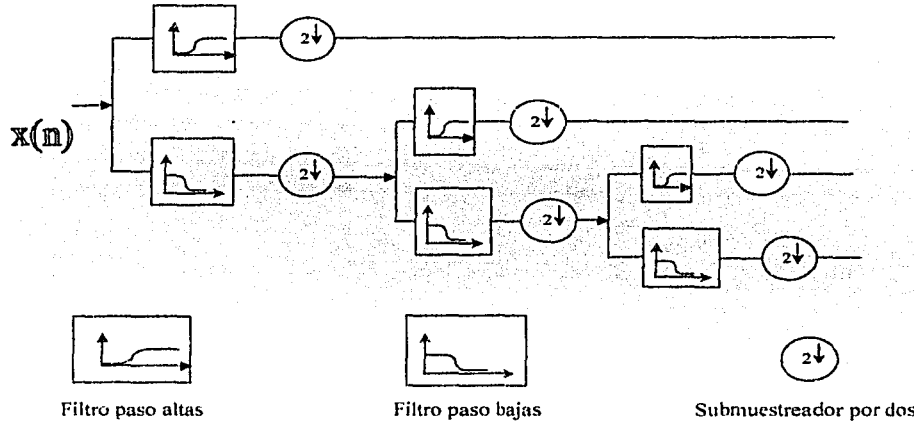


Fig. 2.2 Diagrama de bloques de la DWT implementada a través de una descomposición sub-banda usando un árbol de banco de filtros: El principio de la descomposición wavelet de la señal de entrada $x(n)$ en varias señales de detalle es aplicado. Después de tal descomposición, la rama inferior del diagrama resulta en la señal de aproximación s . Contrario al algoritmo clásico, donde diferentes filtros son usados para obtener las señales individuales de detalle, en este esquema se utilizan los mismos filtros paso altas y paso bajas. Los cambios en la respuesta en frecuencia de los filtros en cada uno de los niveles se logra cambiando la frecuencia de muestreo de la señal, esa es la razón por lo que aparecen bloques de decimación e interpolación en el diagrama. El resultado es que este algoritmo piramidal simplifica considerablemente la implementación e incrementa el desempeño de la transformada wavelet.

Las wavelets usadas como funciones base para la DWT forman un conjunto ortonormal, la cual se define como:

$$\int k_{b,a}(t)k_{b',a'}^*(t)dt = \begin{cases} 1 & \text{si } b = b' \text{ y } a = a' \\ 0 & \text{c.o.c.} \end{cases}$$

Esta propiedad nos garantiza que no existe redundancia en la información obtenida y que cualquier señal finita puede ser representada como una suma ponderada de las funciones base y que la señal puede ser reconstruida perfectamente a partir de un conjunto completo de funciones ponderadas. En general, la escala y el desplazamiento de la familia wavelet son discretizados como:

$$a = a_0^m, b = nb_0 a_0^m; m, n \in \mathbb{Z}$$

Utilizando estos valores, los coeficientes *DWT* de una función continua en el tiempo se definen como:

$$d_{m,n} = \langle f(t), k_{m,n}(t) \rangle = \int f(t) k_{m,n}(t) dt$$

$$f(t) = \sum_m \sum_n d_{m,n} k_{m,n}(t)$$

Dado que esta representación requiere una familia de wavelets de un número finito de funciones, se necesita una función complementaria de naturaleza paso-bajas que permita representar la función usando *L* resoluciones finitas y un componente de aproximación, a la cual llamaremos $\phi(t)$, que representa una pieza clave en el análisis multi-resolución. Fijando los parámetros $a_0=1$ y $b_0=1$, si $k_{m,n}(t) = \psi_{m,n}(t)$, entonces:

$$f(t) = \sum_{n=-\infty}^{\infty} c_{L,n} 2^{-\frac{L}{2}} \phi\left(\frac{t}{2^L} - n\right) + \sum_{m=1}^L \sum_{n=-\infty}^{\infty} d_{m,n} 2^{-\frac{m}{2}} \psi\left(\frac{t}{2^m} - n\right)$$

donde los coeficientes $\{c_{L,n}\}$ son definidos similarmente como

$$c_{L,n} = \langle f(t), \phi_{L,n}(t) \rangle = 2^{-\frac{L}{2}} \int f(t) \phi\left(\frac{t}{2^L} - n\right) dt$$

Las ecuaciones fundamentales de escalamiento son:

$$\psi(t) = 2 \sum_n h_1(n) \phi(2t - n) \quad \text{y} \quad \phi(t) = 2 \sum_n h_0(n) \phi(2t - n)$$

y las condiciones de ortogonalidad que deben cumplir son:

$$\left. \begin{aligned} \int \psi_{j,k}(t) \cdot \psi_{m,n}(t) dt &= \begin{cases} 1 & j = m, k = n \\ 0 & \text{c.o.c.} \end{cases} \\ \int \phi_{j,k}(t) \cdot \phi_{m,n}(t) dt &= \begin{cases} 1 & j = m, k = n \\ 0 & \text{c.o.c.} \end{cases} \\ \int \phi_{j,k}(t) \cdot \psi_{j,k}(t) dt &= 0 \end{aligned} \right\} (2.4)$$

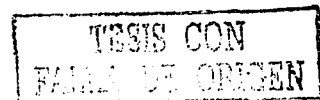
La transformada discreta de Fourier de las secuencias $\{h_0(n)\}$ y $\{h_1(n)\}$ están definidas como:

$$H_0(e^{j\omega}) = \sum_n h_0(n) e^{-j\omega n} \quad \text{y} \quad H_1(e^{j\omega}) = \sum_n h_1(n) e^{-j\omega n} \quad (2.5)$$

De las ecuaciones fundamentales de escalamiento se encuentran las ecuaciones en el dominio de la frecuencia que muestran que es posible generar funciones continuas a través del diseño de funciones discretas:

$$\Psi(\Omega) = H_1(e^{j\omega/2}) \prod_{k=2}^{\infty} H_0(e^{j\omega/2^k}) \quad \text{y} \quad \Phi(\Omega) = \prod_{k=1}^{\infty} H_0(e^{j\omega/2^k}) \quad (2.6)$$

En estas ecuaciones se utilizan frecuencias normalizadas $\omega = \Omega$.



Las condiciones impuestas por las ortonormalidades para lograr una reconstrucción perfecta de la señal original se pueden resumir en [G1]:

$$\sum_n h_0(n)h_0(n+2k) = \delta(k)$$

$$\Rightarrow h_1(n) = (-1)^n h_0(n) \quad (2.7)$$

Por lo tanto, antes de generar las funciones wavelet madre y las funciones de escala, se deben de generar las secuencias $\{h_0(n)\}$ y $\{h_1(n)\}$ ⁸ que cumplan con las condiciones anteriores. A partir de los anterior, se puede afirmar que una base ortonormal wavelet se puede construir a partir de un banco de filtros QMF.

Las principales ventajas que se obtiene con el uso de la DWT en la compresión de imágenes son que existen algoritmos rápidos** para el cálculo de sus coeficientes, se obtiene buen desempeño en la compresión con pérdida y proporciona capacidades de transmisión progresiva. Las siguientes sub-secciones presentan dos de las metodologías que se emplean para la construcción de las familias wavelet.

2.1.5 Análisis multi-resolución [T1]

Partiendo de una torre de espacios $V_k \subset L_2(\mathfrak{R})$ tal que:

1. $V_k \subset V_{k+1}$, para toda k .
2. $\bigcup_k V_k = L_2(\mathfrak{R})$
3. $\bigcap_k V_k = 0$
4. $f(t) \in V_k \Leftrightarrow f(2^{-k}t) \in V_0$
5. Exista una función $\phi \in V_0$ tal que $\{\phi(t-n) : n \text{ entero}\}$ es una base ortonormal de V_0 .

Entonces existe una función $\psi(t) \in V_1$ tal que $\{\psi_{mn}(t) = 2^{m/2} \psi(2^m t - n) : m, n \text{ entero}\}$ que es una base ortonormal para $L_2(\mathfrak{R})$.

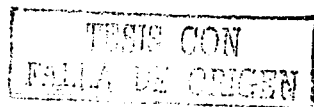
Las primeras tres condiciones aseguran que la torre de sub-espacios del espacio completo de funciones, corresponden a la estructura completa del espacio de la señal. Mientras que en el espacio $L_2(\mathfrak{R})$ toda la textura detallada de las imágenes puede ser representada; en los espacios intermedios V_k , solamente los detalles hasta una resolución fija están disponibles.

La condición cuatro establece que la estructura de los detalles es la misma en cada escala y es únicamente la resolución la que cambia, es decir, una aproximación en una resolución 2^k contienen toda la información necesaria para calcular una aproximación en una resolución más baja 2^{k-1} .

Finalmente la condición cinco implica que el espacio intermedio V_0 y por ende todos los espacios V_k tienen una base simple dada por las traslaciones de una función única.

⁸ En la práctica se emplean 10 u 11 términos para generar Φ y Ψ . [A1]

** El algoritmo en cascada y el esquema lifting.



Hablando en términos de wavelets, podemos decir que las wavelets ortonormales llevan los detalles necesarios para incrementar la resolución de una señal de aproximación, si consideramos el complemento ortogonal W_j^c de V_j que complete la información contenida en V_{j-1} ; la proyección de f sobre V_{j-1} puede ser descompuesta en la suma de las proyecciones ortogonales sobre V_j y W_j^c . De esta manera, el complemento de la proyección de la función sobre W proporciona los detalles que aparecen en la escala $2j$.

Una aproximación multi-resolución es completamente caracterizada por una función de escala ϕ que genera una base ortonormal de cada espacio V_j y cualquier función de escala puede ser especificada a través de un filtro discreto QMF.

Si $2^{-1/2} \phi\left(\frac{t}{2}\right) \in V_1$ por la condición 5, podemos descomponer:

$$2^{-1/2} \phi\left(\frac{t}{2}\right) = \sum_{n=-\infty}^{\infty} h[n] \phi(t-n); \quad h[n] = \left\langle 2^{-1/2} \phi\left(\frac{t}{2}\right), \phi(t-n) \right\rangle$$

donde $h[n]$ es un filtro discreto.

2.1.6 Esquema lifting

El "esquema lifting" (LS) o "lifting" es una técnica para construir intuitivamente bases wavelet en el dominio espacial, o para factorizar filtros wavelet existentes en bloques básicos de construcción, con la finalidad de simplificar su implementación. Su motivación original fue la de construir bases wavelet de segunda generación, es decir, bases que no fueran producto de dilataciones y traslaciones de la wavelet madre. Pero actualmente se ha vuelto popular en sistemas de compresión que requieren; (i) manejo eficiente de la codificación sin pérdida, (ii) uso de memoria mínimo y (iii) baja complejidad computacional.

Daubechies y Sweldens demostraron en [D2] que cualquier transformada wavelet discreta, o filtrado sub-banda de dos bandas con filtros FIR, se podía descomponer en una secuencia finita de simples pasos de filtrado a los que llamaron pasos lifting.^{††} El algoritmo que proponen para obtener este esquema tiene como base la descomposición de la matriz polifase $\tilde{P}(z)$ en matrices elementales y representa una atractiva alternativa al algoritmo clásico de transformación basado en estructura de bancos de filtros.

El uso de esta técnica en las aplicaciones de compresión con pérdida se debe, principalmente, a la eficiencia computacional que se logra; hasta un 100% más alta en el caso de filtros de soporte muy grande [S3]. La ventaja adicional que ha dado auge a este esquema, y que es la causa de que aparezca en el estándar JPEG 2000, es que permite la IWT (Transformada wavelet entera^{†††}) por medio de una modificación directa sobre las funciones wavelet seleccionadas como base. En [A3] se estudia la cuantificación de cualquier degradación en el desempeño asociado con la introducción de propiedades reversibles.

TESIS CON
FALLA DE ORIGEN

^{††} En ocasiones denotadas, estructuras en escalera.

^{†††} También llamada transformación invertible o reversible en aritmética de precisión finita

A continuación se describirá el concepto de este esquema:

El LS calcula la DWT por medio de una secuencia de pasos duales y primales (predicción y actualización de los componentes de las señales). Iniciemos con una señal $x = (x_k)_{k \in \mathbb{Z}}$, que dividimos en dos conjuntos disjuntos llamados componentes polifase:

$$x_e = (x_{2k})_{k \in \mathbb{Z}} \quad \text{muestras pares}$$

$$x_o = (x_{2k+1})_{k \in \mathbb{Z}} \quad \text{muestras impares}$$

Lo que se espera es que estos conjuntos estén altamente correlacionados, por lo que se puede construir un predictor de las muestras pares $P(x_o)$, además de que se debe registrar el error de la predicción:

$$d = x_e - P(x_o)$$

Las muestras pares se obtienen como:

$$x_e = P(x_o) + d$$

Si P es un buen predictor, lo que se espera es que la entropía de primer orden de d sea más pequeña que la de x_o . A la operación de calcular una predicción y registrar el detalle se le llama "paso lifting". Esta idea está relacionada a los esquemas DPCM, y en la aplicación que nos ocupa se completa como sigue; los pasos de predicción pueden tomar en cuenta la correlación espacial, pero para las wavelets es necesario tener alguna información que nos permita discriminar las frecuencias. Con la transformación $(x_e, x_o) \rightarrow (x_e, d)$, por el submuestreo se obtiene aliasing, de hecho el valor promedio de la señal no es el mismo, por lo que es necesario ajustar las muestras pares, para obtener los valores s , a través de un operador de actualización U que se aplica sobre los detalles:

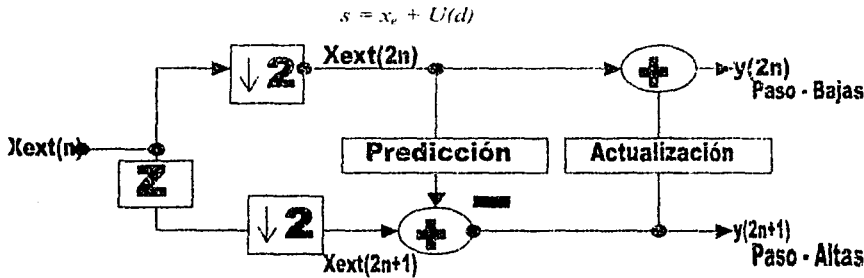


Fig. 2.3 Diagrama de bloques de los pasos lifting de predicción y actualización

Una característica del LS es que no importa como P y U sean seleccionados, el esquema es siempre invertible lo que se traduce en un esquema de banco de filtros de reconstrucción perfecta. La construcción de estas wavelets tiene la desventaja que bajo las mismas condiciones, se pueden diseñar múltiples soluciones, por lo que se debe buscar y seleccionar la más apropiada para la aplicación.

2.1.7 Familias de Wavelets.

Las secciones 2.1.5 y 2.1.6 introdujeron dos técnicas para la construcción de familias wavelets, aunque el análisis de las mismas es tema del capítulo 4, es conveniente presentar su origen para comprender las bases que tome para la inclusión de algunas en el estudio.

La primera característica que fue tomada en cuenta para la construcción de wavelets fueron sus propiedades de localización tiempo-frecuencia, es decir el desvanecimiento de las funciones wavelet en el tiempo y en la frecuencia. Los primeros dos ejemplos presentados muestran esta tendencia:

2.1.7.1 Wavelet Haar

La primera wavelet fue encontrada por Haar [V1]^{*}, la señal original se obtiene a través aproximaciones multi-resolución con funciones constantes. Una de las propiedades de la wavelet de Haar es que tiene soporte compacto, lo que significa que se desvanece fuera de un intervalo finito, de manera inmediata sabemos que esta característica equivale a que no tiene buena localización en frecuencia[†]. La desventaja que presenta es que no es continuamente diferenciable.

La función de escala es: $\phi = 1_{[0,1]}$ (2.9)

El filtro:

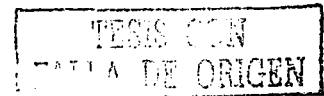
$$h[n] = \begin{cases} 2^{-1/2} & ; n = 0,1 \\ 0 & \text{c.o.c.} \end{cases} \quad (2.10)$$

$$\psi(t) = \begin{cases} -1; & 0 \leq t < 1/2 \\ 1; & 1/2 \leq t < 1 \\ 0; & \text{c.o.c.} \end{cases} \quad (2.11)$$

2.1.7.2 Wavelet Shannon

Se construye a partir de la aproximación multi-resolución de Shannon, la cual aproxima las funciones por su restricción a intervalos de frecuencias bajas. Aunque esta wavelets es infinitamente diferenciable (C^∞) tiene en el tiempo una caída asintótica lenta.

$$\begin{aligned} \phi &= 1_{[-\pi,\pi]} & \hat{h}(\omega) &= \sqrt{2} 1_{[-\pi/2,\pi/2]}(\omega); & \omega &\in [-\pi,\pi] \\ \psi(\omega) &= \begin{cases} \exp(-i\omega/2); & \omega \in [-2\pi,-\pi] \cup [\pi,2\pi] \\ 0 & \text{c.o.c.} \end{cases} \\ \psi(t) &= \frac{\text{sen}2\pi(t-1/2)}{2\pi(t-1/2)} - \frac{\text{sen}\pi(t-1/2)}{\pi(t-1/2)} \end{aligned}$$



2.1.7.3 Wavelets de soporte infinito

En 1985, Y. Meyer construyó las primeras wavelets no triviales, continuamente diferenciables, pero sin soporte compacto (Fig. 2.4). Esta wavelet es una función de banda limitada en frecuencia cuya transformada de Fourier es suave, a diferencia de la transformada de Fourier de la wavelet de

^{*} La primera mención de esta wavelet aparece en el apéndice de la tesis de A. Haar (1909).

[†] Principio de incertidumbre

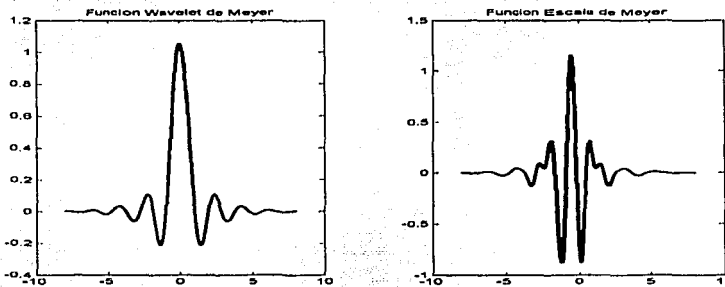
Shannon. Esta suavidad proporciona en el tiempo una caída asintótica mucho más rápida y se construyen a partir de filtros de espejo conjugados $\hat{h}(\omega)$ que son C^n y satisfacen

$$\hat{h}(\omega) = \begin{cases} \sqrt{2}; & \omega \in [-\pi/3, \pi/3] \\ 0; & \omega \in [-\pi, -2\pi/3] \cup [2\pi/3, \pi] \end{cases}$$

$$\hat{\phi}(\omega) = \begin{cases} 2^{-1/2} \hat{h}(\omega/2); & |\omega| \leq 4\pi/3 \\ 0 & |\omega| > 4\pi/3 \end{cases}$$

$$\hat{\psi}(\omega) = \begin{cases} 0; & |\omega| \leq 2\pi/3 \\ 2^{-1/2} \hat{g}(\omega/2) & 2\pi/3 \leq |\omega| \leq 4\pi/3 \\ 2^{-1/2} \exp(-i\omega/2) \hat{h}(\omega/4) & 4\pi/3 \leq |\omega| \leq 8\pi/3 \\ 0; & |\omega| > 8\pi/3 \end{cases}$$

Fig. 2.4



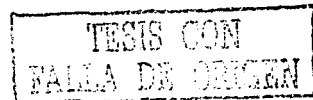
La familia de wavelets gaussianas, también conocida como VMVF (Vanishing Momenta Wavelet Family) se forman con derivadas de gaussianas (Fig. 4.2) y tienen la característica de que no son ortogonales.

Battle y Lemarié en 1988, diseñaron una familia de wavelets que trabajan a partir de aproximaciones multi-resolución con polinomios “spline”, con lo que se puede tener aproximaciones suaves con desvanecimiento asintótico rápido. El espacio V_N de splines de grado $N > 0$, es el conjunto de funciones que son $N-1$ veces diferenciables e iguales a un polinomio de grado N , sobre el intervalo $[N2^j, (N+1)2^j]$ para N entera. En el caso de esta familia ϕ es una B-spline de grado N ,

$$\hat{\phi}(\omega) = (2\pi)^{-1/2} e^{-i\kappa\omega/2} \left(\frac{\text{sen } \omega/2}{\omega/2} \right)^{N+1}$$

$$\text{donde } \kappa = \begin{cases} 0 & \text{para } N \text{ impar} \\ 1 & \text{para } N \text{ par} \end{cases}$$

$$\int \phi(x) dx = 1$$



$$\phi(x) = \begin{cases} 2^{-2M} \sum_{j=0}^{2M+1} \binom{2M+1}{j} \phi(2x - M - 1 + j) & \text{si } N = 2M \text{ es par} \\ 2^{-2M-1} \sum_{j=0}^{2M+2} \binom{2M+2}{j} \phi(2x - M - 1 + j) & \text{si } N = 2M + 1 \text{ es impar} \end{cases} \quad (2.12)$$

Esta wavelet tiene una decadencia exponencial pero con el costo de ser menos regular que las wavelets de Meyer. Para N impar, la wavelet es simétrica con respecto a $1/2$ y la función de escala es simétrica con respecto a 0 , para N par la wavelet es antisimétrica con respecto a $1/2$ y la función de escala es simétrica con respecto a $1/2$.

2.1.7.4 Wavelets ortonormales de soporte finito

El análisis multi-resolución proporcionó el contexto perfecto para la construcción de bases wavelet ortogonales, regulares y de soporte compacto, utilizando este enfoque se podía explicar la existencia de las wavelets de Haar y de Meyer sin que parecieran mágicas. En este punto es importante enfatizar que el enfoque multi-resolución introduce necesariamente la función de escala para permitir el análisis de la función con un número finito de componentes, lo que es primordial en aplicaciones que demanden reconstrucción perfecta. En el caso de la compresión de imágenes, esta condición se hace relevante al seleccionar si la compresión a utilizar es con pérdida o sin pérdida.

Daubechies

En 1988, Ingrid Daubechies construye un conjunto base de funciones wavelet ortonormales, seleccionando los filtros de fase mínima esta selección corresponde a una función wavelet y escala muy asimétrica como se puede ver en la figura 4.11. Posteriormente demostró que la simetría de estas funciones no se puede lograr en bases wavelet ortonormales y que incluso la simetría dependía del valor fijado para:

$$m_0(\omega)^2 = \left| \frac{1}{\sqrt{2}} \sum_n h_n e^{-in\omega} \right|^2 \quad (2.13)$$

En el caso de que se omita la restricción de que la función de escala sea real, entonces la simetría es posible incluso si esta tiene soporte compacto.

En esta familia si N momentos de ϕ son cero $\Rightarrow \phi$ tiene un soporte de tamaño mínimo de $(2N-1)$. Además la regularidad de la función de escala y de la wavelet son iguales.

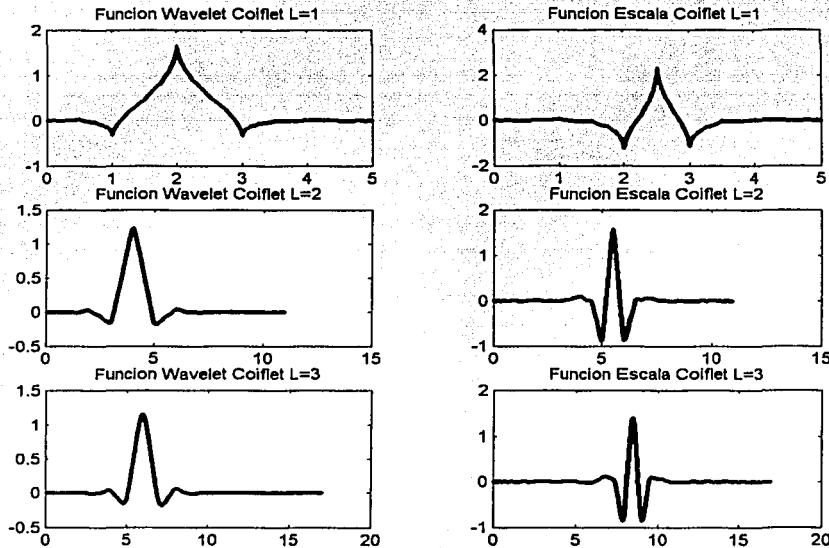
Coiflets

Para aumentar las propiedades de compresión de los filtros de aproximación Coifman sugirió a Daubechies que tanto la función de escala como la de la wavelet tuvieran el mismo número de momentos minimizados, así surgió la familia de Coiflets, donde adicionalmente se cumple con la condición de que:

$$\begin{aligned} \int t^k \phi(t) dt = m_0(k) &= 0 \quad \text{para } k = 1, 2, 3, \dots, L-1 \\ \int t^k \psi(t) dt = m_1(k) &= 0 \quad \text{para } k = 0, 1, 2, 3, \dots, L-1 \end{aligned} \quad (2.14)$$

El momento cero de $m_1(0)$ es requerido para cumplir con las condiciones de ortogonalidad, esto resulta en un momento extra en la función de escala. A este sistema se le llama de grado L y algunas veces tiene el requerimiento adicional de que la longitud N del filtro de la función de escala $h(n)$, sea mínimo.

Fig. 2.5



La propiedad de los momentos en la función wavelet es invariante en el tiempo, pero no sucede lo mismo con los momentos ceros en la función de escala, estos sí dependen del desplazamiento, por lo que es necesario seleccionar apropiadamente este desplazamiento.

Un coiflet con $2K$ momentos minimizados tiene un soporte de tamaño $(6K-1)$ en comparación al tamaño $(4K-1)$ que tendría una wavelet de la familia Daubechies.

2.1.7.5 Wavelets biortogonales

De la esquema filtrado sub-banda, Smith Barnwell demostró que la simetría y la reconstrucción exacta son incompatibles, en el caso de que se utilicen los mismos filtros FIR en la reconstrucción y en la descomposición, sin embargo, si esta condición no es requerida la simetría es posible construyendo un esquema que corresponda a dos bases wavelets duales, asociadas a dos diferentes aproximaciones multi-resolución, a estas familias se les conoce como wavelets biortogonales CDF ya que son el resultados de estudios de Cohen, Daubechies y Feauveau. En esta familia tenemos cuatro filtros en lugar de dos y por ende dos parejas de funciones de escala y wavelet, las condiciones de ortogonalidad son sustituidas por las condiciones de biortogonalidad.

Realizando una descomposición parecida a la que se mostró en la sección 2.1.3, la reconstrucción tiene que ser:

$$a_{m-1,l}(f) = \sum_n [\tilde{h}_{2n-l} a_{m,n}(f) + \tilde{g}_{2n-l} c_{m,n}(f)] \quad (2.15)$$

Donde forzosamente $\tilde{h} \neq h$ y $\tilde{g} \neq g$. La reconstrucción perfecta se logra imponiendo que:

$$\left. \begin{aligned} \tilde{g}_n &= (-1)^n h_{-n+1} \\ \tilde{g}_n &= (-1)^n \tilde{h}_{-n+1} \\ \sum_n h_n \tilde{h}_{n+2k} &= \delta_{k,0} \end{aligned} \right\} \quad (2.16)$$

que equivalen a las condiciones en (2.4).

Si adicionalmente estos filtros cumplen con la condición de que:

$$\prod_{k=1}^{\infty} H(2^{-k} \xi) \quad y \quad \prod_{k=1}^{\infty} H(2^{-k} \xi) \quad (2.17)$$

Se desvanecen más rápido que $C(1+|\xi|)^{-\epsilon-0.5}$ conforme $|\xi| \rightarrow \infty$ para alguna $\epsilon > 0$, donde

$$H(\xi) = 2^{-1/2} \sum_n \tilde{h}_n e^{-jn\xi} \quad y \quad H(\xi) = 2^{-1/2} \sum_n h_n e^{-jn\xi} \quad (2.18)$$

Entonces la función de escala ϕ y su dual $\tilde{\phi}$ se pueden definir como:

$$\phi(x) = \sum_n h_n \phi(2x-n) \quad y \quad \tilde{\phi}(x) = \sum_n \tilde{h}_n \tilde{\phi}(2x-n) \quad (2.19)$$

Y la función wavelet ψ y su dual por:

$$\psi(x) = \sum_n g_n \phi(2x-n) \quad y \quad \tilde{\psi}(x) = \sum_n \tilde{g}_n \tilde{\phi}(2x-n) \quad (2.20)$$

Los coeficientes en la descomposición pueden ser escritos como:

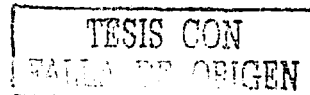
$$\left. \begin{aligned} a_{m,n}(f) &= \langle \phi_{m,n}, f \rangle \\ c_{m,n}(f) &= \langle \psi_{m,n}, f \rangle \end{aligned} \right\} \quad (2.21)$$

$$Y \text{ la reconstrucción es } f = \sum_{m,n} \langle \psi_{m,n}, f \rangle \psi_{m,n} \quad (2.22)$$

Entre más rápido sea el desvanecimiento de los productos en (2.17) $\phi, \tilde{\phi}, \psi$ y $\tilde{\psi}$ serán razonablemente suaves. Las condiciones para que los productos infinitos en (2.17) converjan son:

$$\sum_n h_n = 2^{1/2} \quad y \quad \sum_n \tilde{h}_n = 2^{1/2} \quad (2.23)$$

Adicionalmente, el requisito para que (2.22) se cumpla es:



$$\sum_n (-1)^n h_n = 0 \quad y \quad \sum_n (-1)^n \tilde{h}_n = 0 \quad (2.24)$$

Un hecho importante del uso de wavelets biortogonales es que si ψ tiene \tilde{k} momentos cero, entonces $\langle \psi_{m,n}, f \rangle$ representarán funciones f , \tilde{k} veces diferenciables, con un alto potencial de compresión, ya que muchos coeficientes son lo suficientemente pequeños para ser despreciados [D1].

Si se expresan $H(\xi)$ y $\tilde{H}(\xi)$ como polinomios trigonométricos las condiciones para la reconstrucción perfecta (2.16) para filtros simétricos se reduce a:

$$H(\xi)\tilde{H}(\xi) + H(\xi + \pi)\tilde{H}(\xi + \pi) = 1 \quad (2.25)$$

Buscando que ψ y $\tilde{\psi}$ sean respectivamente $(k-1)$ y $(\tilde{k}-1)$ veces diferenciables, se debe cumplir que:

$$H(\xi)\tilde{H}(\xi) = \cos\left(\frac{\xi}{2}\right)^{2l} \left[\sum_{p=0}^{l-1} \binom{l-1+p}{p} \text{sen}\left(\frac{\xi}{2}\right)^{2p} + \text{sen}\left(\frac{\xi}{2}\right)^{2l} R(\xi) \right] \quad (2.26)$$

Donde $R(\xi)$ es un polinomio impar en $\cos(\xi)$ y $2l = k + \tilde{k}$ lo que implica que la simetría de h y \tilde{h} fuerza a que $k + \tilde{k}$ sea par.

Filtros Spline.

En esta familia $\tilde{\phi}$ es una función B-spline, sus filtros asociados son llamados spline o binomiales para lograr esto en (2.26) $R \equiv 0$ con:

$$\tilde{H}(\xi) = \cos\left(\frac{\xi}{2}\right)^k e^{-j k \xi / 2} \quad \text{donde} \quad \begin{cases} k = 0 \text{ si } \tilde{k} \text{ es par} \\ k = 1 \text{ si } \tilde{k} \text{ es impar} \end{cases} \quad (2.27)$$

Una variante de Spline con longitudes muy similares:

En esta familia $R \equiv 0$ pero el lado derecho de (2.26) se factoriza rompiendo el polinomio de grado $l-1$ en $\text{sen}(\xi/2)$ en un producto de dos polinomios en $\text{sen}(\xi/2)$ con coeficientes reales, uno es asignado a H y otro a \tilde{H} buscando hacer que las longitudes de h y \tilde{h} sean lo más cercanas posible. Los filtros más pequeños de esta familia son $l=4$, $k=4$

2.1.8 Estructuras de descomposición

Wavelet packets

La expansión con wavelet packets fue introducida por Coifman; estas son bases estructuradas en el tiempo discreto o continuo que permiten diferentes relaciones entre la resolución en el tiempo y en la frecuencia. Dado que cada función base tiene una región en el plano tiempo frecuencia, donde la mayoría de su energía es concentrada; la meta es caracterizar la señal con la ayuda de la

transformada wavelet, concentrándose en fenómenos que ocurren en tiempo más pequeños conforme la escala llega a ser más pequeña. Este comportamiento permite una caracterización local de funciones.

La clave del éxito de la transformación sub-banda se encuentra en la selección de la mejor base para la aplicación dada, para esto, el árbol usado debe ser apropiado a las características del espectro de entrada. Una transformación sub-banda puede localizar las características espectrales de una señal por la generación adaptativa del árbol apropiado, en el cual la medida de compactación de energía espectral se emplea para justificar la división de sub-bandas en cada nivel del árbol. Para cada sub-banda W se asigna un costo de información H_w , la cantidad $H_w(S)$ mide que tan caro resulta incluir W en la descomposición usada para representar la imagen S . Definir la mejor base para representar S (con respecto a H_w) es encontrar el subconjunto B_0 que minimiza, para todos los subconjuntos base B :

$$\sum_{w \in B} H_w(S)$$

Cada una de estas cantidades es el número de elementos arriba de un umbral ϵ determinado, es decir:

$$H_w(S) = \#\{x \in S_w, |x| \geq \epsilon\}$$

donde S_w es la secuencia de componentes S que pertenecen a la sub-banda analizada.

El método tradicional para lograr esta división es el "Método de la Mejor Base" [C1], aunque se pueden encontrar otros algoritmos, en [A1] la compactación de energía es calculada para las estructuras de dos y tres niveles, la de mejor compactación se selecciona y el nodo padre es descompuesto en nodos hijos si un cierto umbral es excedido, razón por la cual se le considera a ese nodo significativo; los nodos de las sub-bandas insignificativas

Los siguientes diagramas muestran algunos tipos de descomposición y sus respectivas estructuras de árbol.

Fig. 2.7 Algoritmo de Mallat

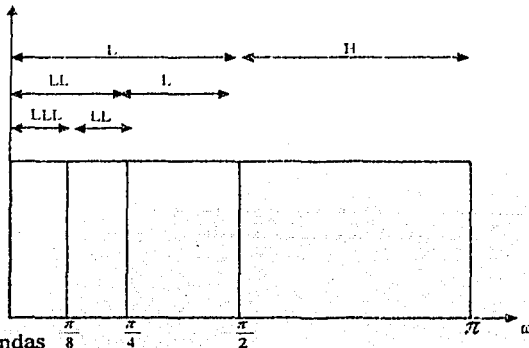
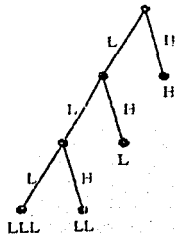
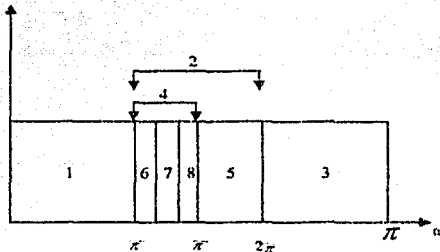
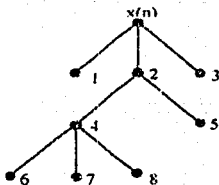


Fig. 2.8 Descomposición en más de dos bandas



2.2 Cuantización

La definición en el diccionario de cuantización es la división de una cantidad en un número discreto de partes pequeñas, con frecuencia suponiéndolas como múltiplos de una cantidad común. En el contexto de la compresión de señales, el propósito de cuantizar la información es la de permitir una representación de precisión finita de los datos.

Aunque se han hecho esfuerzos por emplear la cuantización vectorial en las aplicaciones de compresión, las desventajas que tiene por el entrenamiento requerido - para la formación del diccionario - y la potencia de cómputo - para la búsqueda del óptimo correspondiente - son difíciles de compensar, sin contar que adicionalmente se tiene que almacenar o transmitir el libro de códigos.

Un esquema que ha fructificado en algoritmos modernos es el de cuantizar cada sub-banda de acuerdo a una tasa de bit asignada, posteriormente codificar los coeficientes cuantizados empleando un codificador de entropía. En este ramo algoritmos como el EZW de Shapiro reforzaron la tendencia de los algoritmos de compresión, porque adicionalmente dan como resultado cadenas de código embebidas, requisito para la transmisión progresiva. La mayoría de los métodos tradicionales para codificar coeficientes de transformadas son basados en dos observaciones:

- i) Para la mayoría de las imágenes naturales, la energía está concentrada en la sub-banda de baja frecuencia, lo que permite que el resto de las sub-bandas puedan ser cuantizadas e incluso descartadas sin afectar severamente la calidad de la imagen reconstruida.
- ii) Coeficientes de cada resolución deberían ser independientes a coeficientes de otra resolución. Sin embargo, existe una correlación entre las sub-bandas de las cuales se podría obtener ventaja.

En esta sección se expondrán únicamente los métodos de cuantización incluidos en el JPEG 2000, la siguiente sección proporcionará las bases para la compresión del proceso de codificación de los coeficientes cuantizados.

2.2.1 Escalar

Un cuantizador escalar es una transformación $Q: \mathbb{R} \rightarrow \mathbb{Z}$ que se define a partir de: (1) un conjunto de intervalos o celdas $S = \{S_i; i \in I\}$, donde el conjunto de índices es una colección de enteros consecutivos que inician con 0 ó 1, y (2) un conjunto de valores de reproducción o puntos o niveles $C = \{y_i; i \in I\}$, de tal manera que el cuantizador q se define como $q(x) = y_i$ para $x \in S_i$. Lo anterior, equivale a la siguiente regla de cuantización:

$$q(x) = \sum_i y_i 1_{S_i}(x)$$

donde $1_S(x)$ es 1 si $x \in S$ y 0 de cualquier otra manera. Para esta definición se supone que S es una partición de la línea real, por lo que las celdas son disjuntas y exhaustivas. Un cuantizador es uniforme si los niveles y_i son encuentran equidistantes una cantidad Δ , y los umbrales (límites de los celdas) se encuentran a la mitad de dos niveles adyacentes.

2.2.2 TCQ

Este es un método de baja complejidad que tiene mejor desempeño tasa-distorsión que la cuantización escalar. Está basado en los conceptos de la modulación por códigos trellis, empleando árboles similares a los de la modulación en amplitud de Ungerboeck y la teoría tasa-distorsión de alfabetos restringidos. Algunas de las características de este esquema son:

1. Uso de un libro de códigos de reproducción formado por el conjunto expandido de niveles de cuantización.
2. Basado en una adecuada partición, la estructura del árbol puede reducir el número de niveles de cuantización hasta alcanzar la tasa de codificación deseada.
3. Empleando un libro de códigos determinístico se puede lograr una estructura de codificación computacionalmente simple.
4. El codificador usa el algoritmo de Viterbi y, en la mayoría de los casos, la codificación requiere únicamente, cuatro multiplicaciones, cuatro sumas y cuatro cuantizaciones escalares por muestra de fuentes mas dos sumas y una comparación por estado de trellis por muestra de la fuente.
5. Un análisis de los efectos de los errores del canal muestra que el TCQ puede ser construido para ser insensitivo a los errores de canal.[M2]

En cuanto a la teoría tasa-distorsión para una fuente discreta, un alfabeto de reproducción específico debe ser seleccionado para calcular la función de distorsión; la idea es encontrar una expresión para lograr el mejor performance posible al codificar una fuente continua usando un alfabeto de reproducción finita. La opciones disponibles cuando se selecciona un alfabeto de salida son:

- a) Seleccionar solamente el tamaño del alfabeto (número de elementos)
- b) Seleccionar el tamaño y los valores reales del alfabeto.
- c) Seleccionar el tamaño, valores y las probabilidades con las cuales los valores serán usados.

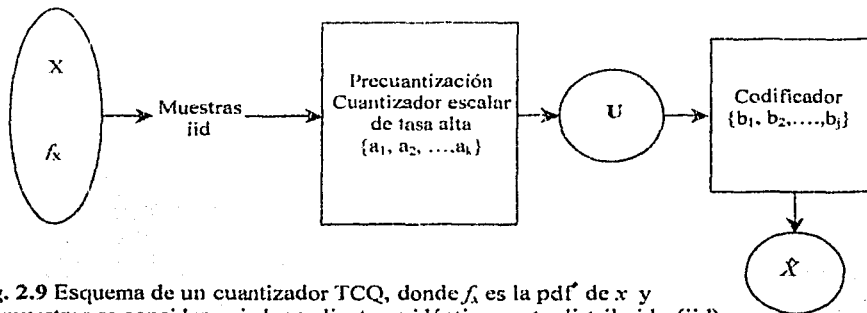


Fig. 2.9 Esquema de un cuantizador TCQ, donde f_x es la pdf^{*} de x y las muestras se consideran independientes e idénticamente distribuidas(iid)

De la figura 2.9, el ruido de cuantización se obtiene como $Q=U-X$ y la distorsión total del sistema considerando que se trata de un cuantizador Lloyd-Max es

$$E[(X - \hat{X})^2] = E[(X - U)^2] + E[(U - \hat{X})^2]$$

* Función de distribución de probabilidad

Usando el algoritmo de Blahut para fuentes discretas se calcula la función tasa distorsión $D_0(R)$ para la fuente discreta U y un alfabeto de reproducción $\{b_1, b_2, \dots, b_j\}$. La función de distorsión del sistema es:

$$D_c(R) = E[(X - U)^2] + D_0(R)$$

El codificador para X se construye poniendo en cascada un cuantizador Lloyd Max con un codificador para U que use símbolos $\{b_1, b_2, \dots, b_j\}$ mientras logra un performance arbitrariamente cercano a $D_0(R)$.

En cuanto a la teoría de modulación, para transmitir uno de 2^m símbolos por un intervalo de la señal, la constelación de puntos de 2^m es ampliada a 2^{m+1} puntos y particionada en $2^{\hat{m}+1}$ subconjuntos donde \hat{m} es un entero menor o igual a m . \hat{m} de los bits de entrada son expandidos por un código convolucional a una tasa $\hat{m}/(\hat{m}+1)$ y usados para seleccionar de cual subconjunto, el símbolo de canal proviene. El resto de los $m - \hat{m}$ bits son usados para seleccionar uno de los $2^{m-\hat{m}}$ símbolos de canal en el subconjunto seleccionado. Como se supone que el canal corrompió la señal con ruido, la decodificación de Viterbi es usado para encontrar la secuencia de símbolos que minimizan la probabilidad que las secuencias transmitida y recibida sean diferentes.

Este esquema mejora las técnica de modulación porque el código convolucional y la partición en conjuntos son seleccionados para incrementar la distancia euclidiana entre las secuencias permitidas de los símbolos de canal. El algoritmo de Viterbi es usado para encontrar la secuencia disponible de símbolos de canal que es más cercana, en cuanto a distancia euclidiana, a la secuencia recibida a la salida del canal.

La distancia euclidiana entre dos secuencias de longitud n (x y \hat{x}) es:

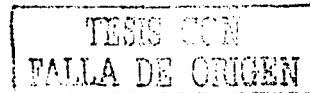
$$d_E(x, \hat{x}) = \sqrt{\sum_{i=1}^n (x_i - \hat{x}_i)^2}$$

El problema es: dada x , encontrar la secuencia \hat{x} que minimiza $d_E(x, \hat{x})$, lo que equivale a minimizar $\left(\frac{1}{n}\right) d_E^2(x, \hat{x}) = \rho_n(x, \hat{x})$ que es la distorsión del error cuadrático.

Ya que cualquier conjunto de secuencias $C = \{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_k\}$, cada una de longitud n , define un código de fuente; el conjunto de todas las secuencias de canal permitidas y el decodificador de Viterbi pueden ser usados como un código fuente y su correspondiente codificador de fuente. Específicamente, dada una secuencia de x , el algoritmo de Viterbi encuentra la secuencia de C que minimiza $\rho_n(x, \hat{x})$.

2.3 Codificación por entropía

A continuación se presentan brevemente los algoritmos que representan el antecedente al codificador por entropía del JPEG 2000 y del JBIG2, la tendencia que se puede observar es la codificación de planos de bits, codificando los bits del MSB al LSB y considerando la información de signo como un plano adicional.



2.3.1 Codificador MQ

Es un codificador aritmético de detección binaria basado en los contextos detectados, su antecesor es el codificador Q [P1]. La sub-optimización del codificador es compensada por una sofisticada estimación de la probabilidad empleando autómatas, lo cual le da como ventaja su rápida adaptación a los datos de la fuente, ya que en lugar de mantener un contador de píxeles, el proceso de estimación es implementado como un autómata de 94 estados.

El autómata es una cadena de Markov que contiene un estado por cada estimado de probabilidad, los estados son organizados en renglones ordenados por un nivel de adaptación. El modelo consiste en apuntadores a los estados del autómata, uno por cada contexto. El codificador y decodificador estiman el modelo dinámicamente en la codificación y decodificación, respectivamente, manejando la estimación de la probabilidad por estimación.

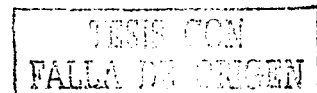
El proceso de adaptación inicia en el estado 0 con un modelo de probabilidades LPS, integrado al codificador, encuentran espaciadas uniformemente. En un punto dado del proceso, la cadena de símbolos es mapeada a un sub-intervalo único $[c, c+a) \subseteq [0,1)$ donde $c=C \cdot 2^{-16 \cdot n}$, $a=A \cdot 2^{-16 \cdot N}$ y $C, A \in \mathbb{Z}$. En este marco N es el número total de desplazamientos en la normalización que aseguran que $2^{15} \leq A < 2^{16}$. La cadena de bits comprimida es representada por C y su longitud puede ser $N+2$ o menos bits dependiendo de las reglas de terminación del codificador y decodificador.

Supongamos que el siguiente símbolo ocurre en el contexto k , en el cual el LPS (Símbolo Menos Probable) tiene una estimación de probabilidad $p_k \in (0, 1/2)$. Idealmente la longitud del intervalo se contrae de acuerdo a que $a \leftarrow a \cdot p_k$, si el siguiente símbolo es el LPS y $a \leftarrow a \cdot \rho_k$, si el siguiente símbolo es el MPS (Símbolo Más Probable). Es posible aproximar $A \leftarrow \rho_k$ en el primer caso y $A \leftarrow A - \rho_k$ en el segundo, donde ρ_k es la aproximación entera de $p_k \alpha 2^{16}$ con $\alpha \cong 0.7^*$. A continuación el símbolo LPS es mapeado al sub-intervalo más bajo tal que C no sufra cambios. Cuando un MPS ocurre $C \leftarrow C + \rho_k$.

Si en este proceso $A < 2^{15}$, se ejecuta una re-normalización para ajustar el rango de A , durante este proceso A y C son duplicados y N es modificado de acuerdo a estos incrementos, hasta que A es restaurado en su rango original. Para probabilidades del LPS cercanos a $1/2$, el procedimiento es modificado por el mecanismo de intercambio condicional con la finalidad de mejorar la aproximación del intervalo. La política de normalización asegura que A pueda ser representado por 16 bits, sin embargo, en cualquier punto del proceso C tiene una representación de $N+16$ bits. Todos las operaciones aritméticas toman lugar en los 16 bits menos significativos de C , pero se debe estar prevenido para enviar los bits más significativos a un buffer de salida.

El codificador MQ incorpora un mecanismo de llenado de bits, que permite la liberación de bytes de código con un costo de implementación menor que en el codificador Q y que permite almacenar los bits consistentes hasta que el acarreo pueda ser resuelto. Los bytes reconstruidos son liberados a través de un buffer de sólo un byte, siempre que este buffer asume el valor FF, un bit extra es insertado en la representación de C evitando que el acarreo se pueda propagar más allá del único byte almacenado en este buffer. El llenado de bits añade aproximadamente 0.05% de la longitud de código y tiene la característica de que cualquier pareja de bytes liberados a la cadena de datos comprimidos caen en el rango de 0 a FF8F.

* Este dato es una media empírica de $2^{16}A$.



2.3.2 EZW (Embedded Zerotree Wavelet).

Desde su presentación por Shapiro en [S2], los algoritmos de codificación que se basan en wavelets presentan estructuras que sacan provecho no sólo de la correlación que existe en los coeficientes de cada una de las bandas, sino de la correlación entre sub-bandas. El zerotree es una estructura de datos que permite representar un fenómeno que ocurre en los coeficientes wavelet: cuando un coeficiente es irrelevante a una resolución dada, es irrelevante en resoluciones más bajas, dando como resultado una estructura de árbol de ceros.

El objetivo del EZW es la codificación embebida de los coeficientes wavelet empleando zerotree, a partir de una secuencia de decisiones binarias que distinguen una imagen dada de la imagen nula. Los conceptos en los que se basa el algoritmo son:

- ❑ Una transformada wavelet discreta u otra descomposición jerárquica de sub-banda que permita una representación multi-resolución compacta.
- ❑ Predicción de la ausencia de información significativa a través de las escalas de la imagen, esto se realiza explotando la autocorrelación.
- ❑ Codificación zerotree para una representación eficiente del mapa de significado multi-resolución.⁴
- ❑ Cuantización de aproximación sucesiva que proporciona una representación compacta multi-precisión de los coeficientes significativos, lo que facilita la característica de cadena embebida.
- ❑ Codificación aritmética adaptiva que no requiere tablas pre-almacenadas o entrenamiento alguno para lograr una compresión sin pérdida.
- ❑ Cálculo de la tasa de decodificación y la función tasa-distorsión punto a punto.

Los problemas en los que se centra la solución expuesta son:

- > Obtener la mejor calidad de la imagen para una tasa de bit dada.
- > Obtener una cadena embebida.

Componentes:

1) Transformada Wavelet Discreta (DWT) en un esquema de sub-banda jerárquico.

Aunque la imagen ha sido transformada usando una transformación que des-correlaciona, las ocurrencias de los coeficientes insignificativos no son eventos independientes.⁵

2) Predicción de la ausencia de información significativa a través de las escalas explotando la similitud inherente en las imágenes.

Los "zerotrees" permite la predicción exitosa de coeficientes insignificativos a través de las escalas, las cuales pueden ser representadas eficientemente como árboles con crecimiento exponencial. Los símbolos que se emplean en la codificación del mapa de significación son:

- a) Raíz zerotree (No se usa en la escala más fina)
- b) Cero aislado
- c) Significante negativo

⁴Mapas que indican las posiciones de los coeficientes significativos.

⁵El cálculo del umbral inicial T_0 se basa en esta dependencia.

d) Significante

3) Cuantización por aproximación sucesiva.

Se emplea un protocolo de prioridades por medio del cual el ordenamiento por importancia es determinado por la precisión, magnitud, escala y localización espacial de los coeficientes waveletes. (En particular, los coeficientes más grandes en magnitud son considerados más importantes que los más pequeños, no importando su escala).

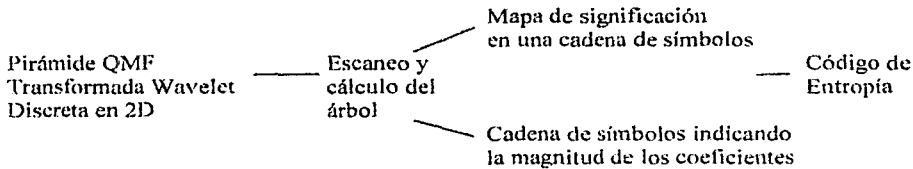
4) Compresión universal sin pérdida a través de una codificación aritmética multi-nivel adaptiva.

Proporciona un método rápido y eficiente para codificar la entropía de cadenas de símbolos.

El algoritmo se ejecuta secuencialmente y se detiene cuando se cumple una de las siguientes condiciones:

- a) Una tasa de bit en el destino.
- b) Una distorsión en el destino.

Se puede encontrar exactamente la tasa de bit que permite el destino y calcular una función de la tasa operacional contra la distorsión punto a punto.



Los coeficientes de la transformada son estadísticamente independientes y se pueden modelar como variables aleatorias con media nula, aunque quizá no distribuidas idénticamente. Las pdf's (funciones de distribución de probabilidad) de los coeficientes son simétricos.

La entropía de la distribución resultante de los índices de los coeficientes (índice de marca=BI) debe ser lo suficientemente pequeña para que los símbolos puedan ser codificados por entropía en una tasa de bit destino baja. El medio paso del cuantizador debe ser simétrico, quizá no uniforme y diferente para diferentes grupos de coeficientes transformados. Si el BI central tiene índice 0, debido a la simetría, para un BI diferente de cero, un índice positivo o negativo es igualmente probable, por lo que se requiere un bit para el signo.

Un código de entropía puede diseñarse basado en el modelado de las probabilidades de los BI como la parte de los coeficientes que tienen el mismo valor absoluto del índice en particular. Suponiendo que los símbolos resultantes son independientes, la entropía H de los símbolos será:

$$H = -p \log_2 p - (1-p) \log_2 (1-p) + (1-p) [1 + H_{Nz}]$$

Probabilidad de que el coeficiente sea cuantizado a 0

Entropía binaria del mapa de significación

Entropía condicional de la distribución de los valores no nulos

Entropía condicional del valor absoluto de los coeficientes cuantizados dado que no son cero.

Si manejamos H como una medida del costo de codificación, la expresión anterior puede ser reescrita:

$$\text{Costo Total} = \text{Costo del mapa de significación} + \text{Costo de los valores no nulos.}$$

Se dice que x es insignificante con respecto al umbral T , si x es un coeficiente wavelet donde $|x| < T$

Si un coeficiente wavelet en una escala ordinaria es insignificante con respecto a un umbral T , entonces todos los coeficientes wavelet de la misma orientación y en la misma posición espacial en escalas más finas son, con un alto grado de probabilidad, insignificantes con respecto a T .

En un sistema de sub-banda jerárquico, cada coeficiente en una escala dada intermedia puede ser relacionado a un conjunto de coeficientes en la siguiente escala más fina de la misma orientación: El coeficiente en la escala grande es llamado padre, y todos los coeficientes que corresponden a la misma posición espacial en una escala más fina son llamados hijos; y todos los correspondientes en todas las escalas más finas de la misma orientación son llamados descendientes. Similarmente, para un hijo dado, el conjunto de todos los coeficientes en todas las escalas más burdas de la misma orientación y posición son llamados ancestros.

Con la excepción de las sub-bandas de frecuencia más baja, todos los padres tienen cuatro hijos. Para la sub-banda de frecuencia más baja, la relación padre-hijo se define de tal manera que cada nodo padre tiene tres hijos.

Para formar el mapa de significación, un reconocimiento de los coeficientes se realiza de tal manera que ningún nodo hijo es analizado antes que a su padre. Para una transformación en N escalas, el barrido comienza en la sub-banda de frecuencia más baja denotada por LL_N y se examinan las sub-bandas HL_N , LH_N y HH_N , posteriormente se mueve a la escala $N-1$. Se considera x como un elemento del zerotree, si el mismo y todos sus descendientes son insignificantes con respecto a T . Un elemento de un zerotree para un umbral T es una raíz zerotree si no es descendiente de una raíz zerotree previamente encontrada, es decir, no es un elemento insignificante predecible a partir del descubrimiento de una raíz zerotree en una escala mayor considerando el mismo umbral T . La raíz zerotree indica que la detección de insignificantes entre los coeficientes en escalas más finas es completamente predecible. Finalmente, el mapa de significación puede ser representado eficientemente como una cadena de símbolos.

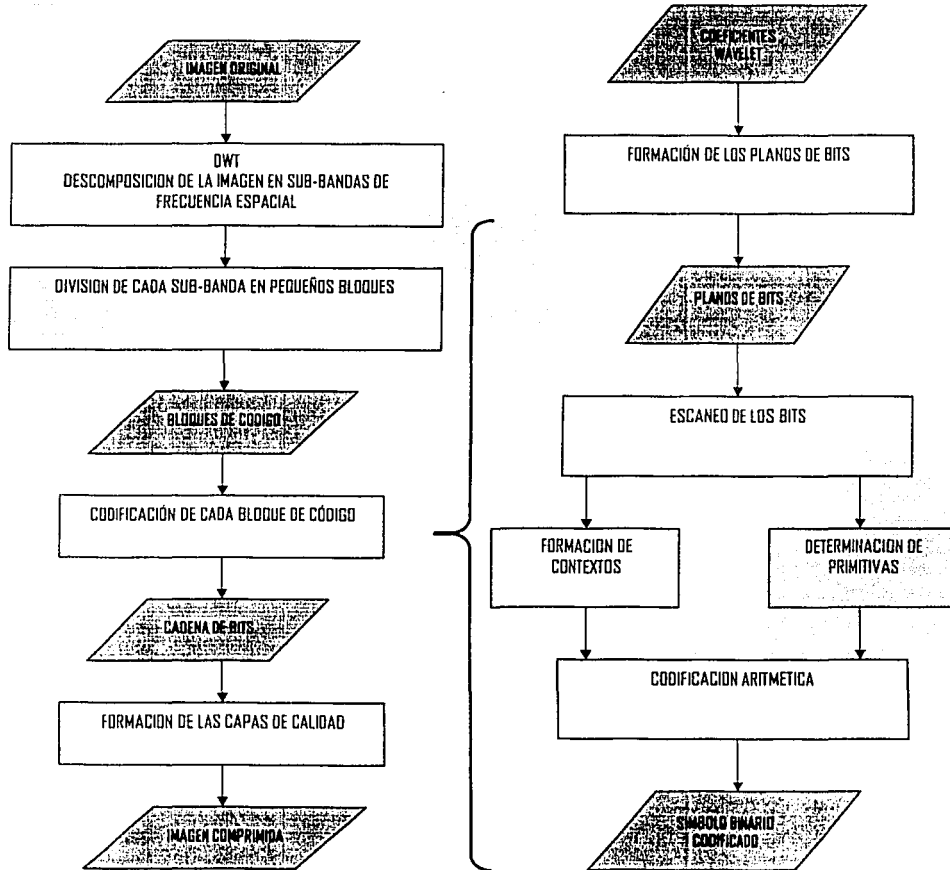
2.3.3 Codificación embebida de bloques.

En este rubro podemos encontrar algoritmo como el SPIHT [S7] con codificación aritmética y el EBCOT [T2]. El punto importante en estos algoritmos es el diseño cuidadoso de las primitivas de codificación de plano de bits. Las ventajas de este esquema son:

- Localización en memoria.
- Acceso aleatorio espacial a la imagen.
- Facilidad de manipulación geométrica
- La cadena de bits generada puede ser truncada en cualquiera de un número grande de puntos.
- El compresor es libre para asignar contribuciones adicionales a cada capa de calidad de un bloque en cualquier etapa de la codificación. Un enfoque es calcular la información de tasa y distorsión (R_n , D_n) de cada punto de truncamiento candidato n , en la cadena embebida generada para el bloque B_n , y usar esta información para asignar las contribuciones del

bloque a las capas de calidad de tal forma que, las capas 1,2,...,λ contengan una representación óptima en el sentido tasa-distorsión para cada λ.

Fig 2.10 Diagrama del esquema general del EBCOT.



Codificación de los planos de bits.

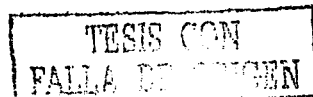
Sea $\chi[n] \in \{1, -1\}$ el signo de la muestra** $s[n]$ en la localización $n=[n_1, n_2]$ y $v[n]$ la magnitud cuantizada de $s[n]$ donde δ es el tamaño de paso de cuantización del intervalo en cero. Entonces:

$$v[n] = \left\lfloor \frac{s[n]}{\delta} \right\rfloor$$

El valor de cualquier muestra individual $s[k]$ es representado por la codificación progresiva de sus bits del MSB al LSB. Si p es el número de planos de bits para representar todas las muestras de un bloque, entonces $v[k] < 2^p \forall k \in B$, y $v_p[k]$, con $p = (P-1) \dots 0$, es el valor del p -ésimo bit de magnitud. En todos los casos, el signo $\chi[k]$ es codificado inmediatamente después de que se codifica la posición del primer bit significativo (es decir, después de que se codifica el primer $v_p[k]=i$).

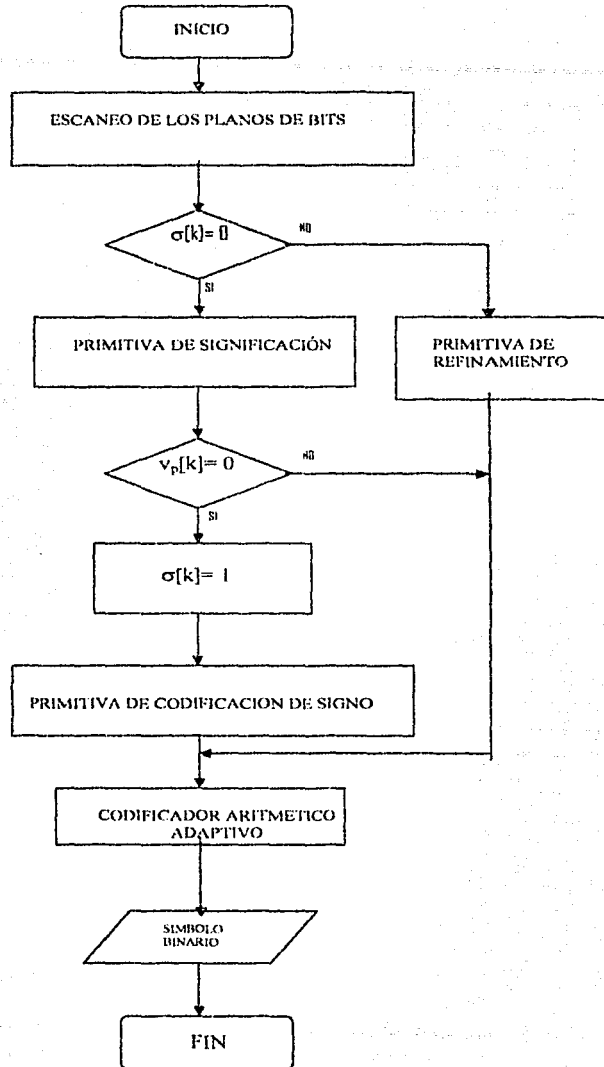
El truncamiento de la cadena embebida equivale a descartar un número p_i de bits de magnitud significativa de cada muestra $s[k]$; esto equivale a emplear un paso de cuantización más grande e igual a $2^{p_i} \delta$, para esa muestra. El codificador utiliza un modelo basado en contextos para determinar tanto el orden como la codificación de los eventos.

Primitivas utilizadas en la codificación de los planos de los bits.- Las primitivas son usadas para codificar nueva información de algún plano p de la muestra $s[k]$, obviamente los bits de magnitud más significativos ya fueron codificados en ese momento. $\sigma[k]$ es una variable binaria de estado que indica si un bit de magnitud con valor distinto de cero ya ha sido codificado para $s[k]$. $\sigma[k]=1$ si $v[k] \geq 2^{p-1}$. Si $\sigma[k] = 0$ (la muestra todavía no es significativa) una primitiva de "Codificación de significación" es usada para codificar si el símbolo llega a ser o no significativo en el plano p ($v_p[k]=1$). Si esto ocurre el estado de $\sigma[k]$ cambia a 1 y la primitiva de "Codificación de signo" debe también ser invocada para identificar el signo $\chi[k]$. Por el contrario si $\sigma[k] = 1$, la muestra ya es significativa, y la primitiva de "Refinamiento" es usada para codificar el valor de $v_p[k]$. Este algoritmo se presenta en la figura 2.11.



** En realidad, podrían tratarse de muestras, coeficientes de la transformada o incluso estos mismos coeficientes pero cuantizados.

Fig 2.11 Uso de las primitivas



Capítulo 3

JPEG 2000

Este capítulo pretende exponer el contexto en el que se aplicarán las wavelets seleccionadas, por lo que se centra en la descripción del estándar JPEG 2000 y principalmente a la parte I del mismo que representa su esencia. La parte I llegó a ser estándar en diciembre de 2000 (ITU-T.800 e IS 15444).

3.1 Objetivo del estándar.

Crear un estándar para un sistema de compresión capaz de aplicarse a varios tipos diferentes de imágenes estáticas (binivel, de niveles de gris, de color) con características diferentes (imágenes naturales, científicas, médicas, militares, de texto y que contengan gráficos) preferentemente dentro de un sistema unificado.

3.2 Partes del JPEG 2000.

Parte	Nombre	Observaciones
I	Sistema de Codificación de Imágenes JPEG 2000	Núcleo de la aplicación, intenta limitar el número de posibles algoritmos de codificación para proporcionar el máximo intercambio posible.
II	Extensiones	Añade más propiedades y sofisticación al núcleo, consiste de tecnologías opcionales de "valor agregado", no requeridas en todas las implementaciones.
III	Motion JPEG 2000	Es un método usado para editar video de alta calidad sin la existencia de un estándar ISO
IV	Conformación de pruebas	Conjunto de pruebas que aseguran las implementaciones de calidad del estándar.
V	Software de referencia	Grupo J2000 de Cannon France, Eriesson, EPFL Jasper de University of British Columbia
VI	Formato de archivos para imágenes compuestas	Para aplicaciones como fax y pre-impresión
VII	Reporte técnico	

3.3 Características generales.

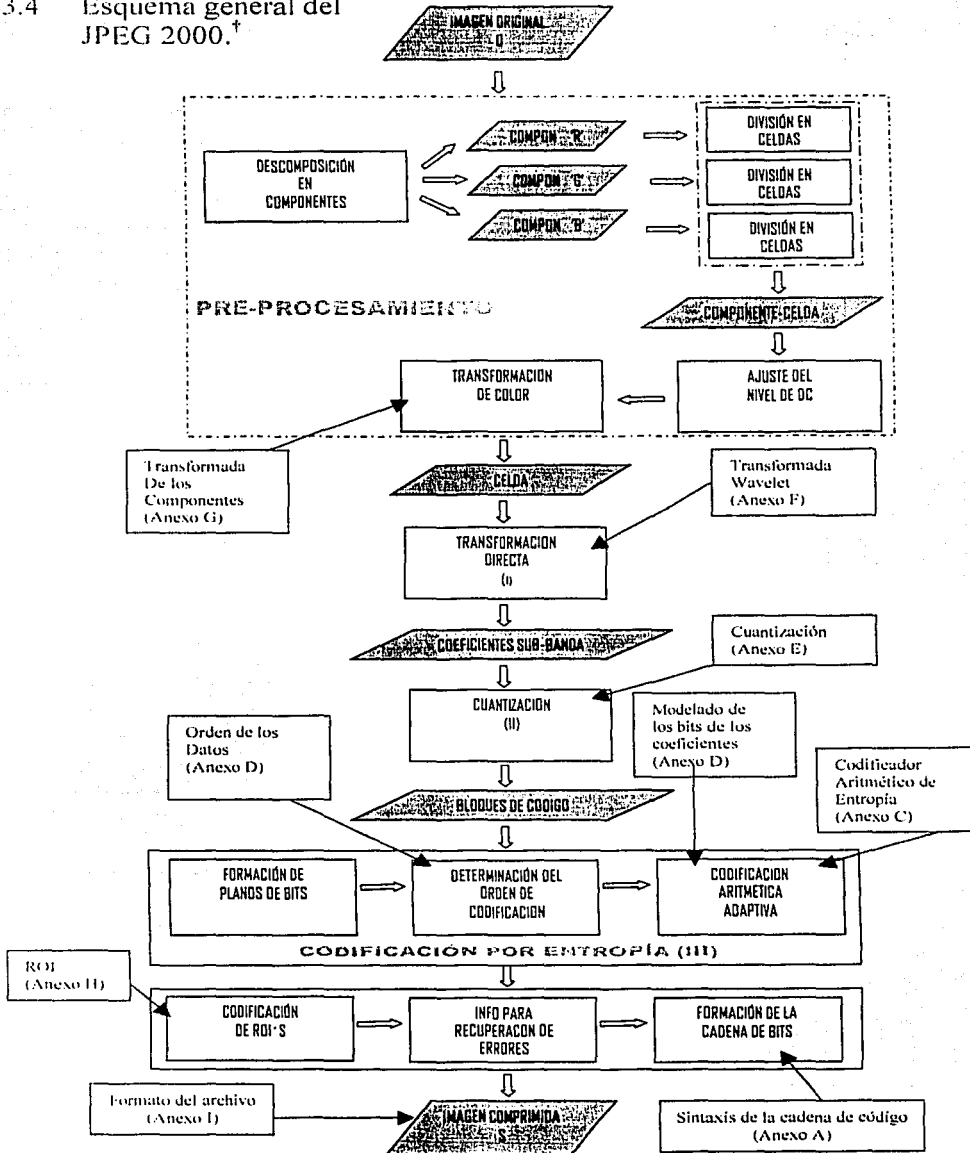
1. Ofrece un desempeño superior a bajas tasas de bits (es decir, a 0.25 bpp) sin que esto implique un sacrificio en la optimización de la función R-D (tasa-distorsión).
2. Comprime y descomprime imágenes con varios rangos dinámicos (es decir, 1 a 6 bits) por cada componente de color
3. Proporciona compresión sin pérdida en el curso de la decodificación progresiva.
4. Apropiado para la transmisión en tiempo real.
5. Transmisión progresiva que permita la escalabilidad de resolución espacial y de calidad (SNR).
6. Considera criterios donde involucra la calidad subjetiva de la imagen considerando el sistema visual humano.
7. Codificación de ROI's (Regiones de interés) en una calidad mejor y menor distorsión que el resto de la imagen
8. Permitir una arquitectura abierta* para optimizar el sistema a los diferentes tipos de imágenes y aplicaciones.
9. Diseño apropiado de la cadena codificada que puede ayudar a los subsecuentes sistemas de corrección de errores en el decremento de las fallas catastróficas de decodificación. Apropiado para el diseño del codificador de canal.
10. Protección a la seguridad de las imágenes. Facilidades para añadir información de derechos de autor y uso.

Estas funcionalidades son posibles dadas las etapas que sigue el estándar. Éstas son esquematizadas a continuación:

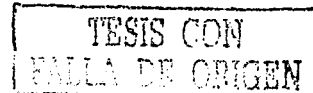
* Arquitecturas cliente / servidor, por ejemplo.

TESIS CON
FALLA DE ORIGEN

3.4 Esquema general del JPEG 2000.[†]



[†] Se marcan los anexos donde se describen los bloques en el estándar JPEG 2000



3.5 Arquitectura.

Pre-procesamiento

1. División en celdas

Se refiere a la partición de la imagen original en bloques rectangulares no sobrepuestos (celdas), las cuales son comprimidas independientemente, como si las celdas fueran imágenes distintas.

Esta operación reduce los requerimientos de memoria y pueden ser usadas para decodificar partes específicas de la imagen ya que las celdas pueden ser reconstruidas independientemente.

Consideraciones

- 1.1 Los componentes de donde provienen las celdas no necesariamente tienen que tener el mismo número de bits por muestra, ni necesitan ser todos signados o no signados.
- 1.2 Todas las celdas tienen exactamente las mismas dimensiones, excepto si es el caso, en los límites de la imagen.
- 1.3 El tamaño de las celdas es arbitrario incluso puede ocuparse la imagen completa.
- 1.4 Los componentes pueden tener diferentes valores de sub-muestreo, aunque es requerido posteriormente ajustar la imagen completa a una rejilla de alta resolución que asegure la consistencia espacial en las celdas resultantes de los componentes.
- 1.5 El uso de este proceso afecta tanto la calidad subjetiva como objetiva de la imagen.
Usar celdas pequeñas crean más artefactos de celda que usar celdas más grandes. La degradación de la imagen es más severa en el caso de tasas de bits bajas.

2. Ajuste de DC

Este proceso sustrae la misma cantidad 2^p , donde p es la precisión en bits del componente. En el decodificador esta misma cantidad es sumada.

Consideraciones

- 2.1 El desplazamiento de DC es realizado sobre muestras de los componentes que son solamente no signados.
- 2.2 El desplazamiento del nivel no afecta la varianza
- 2.3 El desplazamiento convierte una representación no signada a una representación complemento a 2

3. Transformación de los componentes. Transformación de color

Proporciona el medio para eliminar la correlación entre los componentes, mejorando la compresión y permitiendo la cuantización relevante visualmente

Consideraciones

3.1 Se soportan dos transformaciones de los componentes;

- (i) Transformación irreversible (ICT), para la codificación con pérdida

(ii) Transformación reversible (RCT) que puede ser usada para codificación con pérdida y sin pérdida.

3.2 Es posible no emplear ninguna transformación

3.3 Si se emplea ICT[‡] se debe forzosamente usar la transformación wavelet irreversible 9/7.

Las fórmulas que emplea ICT son :

$$\begin{bmatrix} \gamma \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.16875 & -0.33126 & 0.5 \\ 0.5 & -0.41869 & -0.08131 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \text{TRANSFORMACIÓN DIRECTA}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.0 & 0 & 1.402 \\ 1.0 & -0.34413 & -0.71414 \\ 1.0 & 1.772 & 0 \end{bmatrix} \begin{bmatrix} \gamma \\ C_b \\ C_r \end{bmatrix} \quad \text{TRANSFORMACIÓN INVERSA}$$

3.4 Ya que la RCT está definida de enteros a enteros, se puede usar para la codificación con pérdida y sin pérdida, aunque sólo en combinación con la transformación wavelet reversible. Esta es una transformación de descorrelación, por lo que también es llamada descorrelación del color, que se aplica a las tres primeras componentes de una imagen. Sus metas son:

- Lograr una compresión eficiente
- Lograr un espacio de color razonable para la cuantización con respecto al sistema visual humano
- Permitir la posibilidad de obtener una descompresión sin pérdida (reconstrucción exacta con una precisión entera finita)

Dado que esta es una aproximación a la transformación de YUV, se deben cumplir las siguientes restricciones:

- Todos los componentes deberán tener los mismo parámetros de muestreo
- Todos los componentes deberán tener el mismo número de bits por muestra
- Deben existir al menos tres componentes

Las fórmulas que se emplean son:

$$\begin{bmatrix} \gamma_r \\ v_r \\ u_r \end{bmatrix} = \begin{bmatrix} \lfloor \frac{R+2G+B}{4} \rfloor \\ R-G \\ B-G \end{bmatrix} \quad \begin{array}{l} \text{El subíndice } r \text{ significa reversible y,} \\ \lfloor a \rfloor \text{ indica el entero más grande que no excede a } a \end{array}$$

TESIS CON
FALLA DE ORIGEN

[‡] Una forma efectiva de reducir los datos de los componentes es usar ICT seguido por un submuestreo de la crominancia (C_r, C_b) [J1], lo que se presenta de manera natural en el estándar JPEG 2000 al usar la transformación wavelet.

$$\begin{bmatrix} G \\ R \\ B \end{bmatrix} = \begin{bmatrix} \gamma_r - \left[\frac{U_r + V_r}{4} \right] \\ V_r + G \\ U_r + G \end{bmatrix}$$

Procesamiento principal

4. Transformada wavelet directa

Consideraciones de la DWT:

- 4.1 En la Parte I del JPEG 2000 solamente descomposiciones potencias de 2 son permitidas en forma de descomposiciones "diádicas".
- 4.2 El estándar usa una descomposición en sub-bandas de una dimensión de un conjunto de muestras de 1D en muestras paso-bajas y paso-altas. Las muestras paso bajas representan una versión submuestreada y de baja resolución del conjunto original. Las muestras paso-altas representan una versión residual submuestreada del conjunto original, necesaria para la reconstrucción perfecta del conjunto original a partir del conjunto paso-bajas. La DWT puede ser reversible o irreversible.
- 4.3 La transformada irreversible, que proporciona la compresión más alta, es implementada por medio del filtro 9/7.
- 4.4 La transformada reversible, cuya complejidad es menor a la transformada irreversible, es implementada por medio del filtro 5/3 del Le Gall y permite la compresión sin pérdida.
- 4.5 El estándar soporta dos modos de filtrado basado en convolución y basado en lifting.
- 4.6 No importando el tipo de modo de filtrado, la señal debe ser extendida periódicamente para asegurar que por las operaciones de filtrado que toman lugar en ambos límites de la señal, una muestra de la señal exista y corresponda espacialmente a cada coeficiente de la máscara del filtro. El número de muestras adicionales depende de la longitud del filtro. La extensión simétrica de los límites es del tipo (1,1) - la 1ª y última muestra aparece solamente una vez y son muestras importantes (WS); la longitud del kernel es impar.
- 4.7 El filtrado por convolución consiste en realizar una serie de productos puntos entre las máscaras de los filtros y la señal extendida.
- 4.8 El filtrado basado en lifting consiste en una secuencia de operaciones de filtrado simples para los cuales se actualizan alternándose los valores de las muestras impares con sumas ponderadas de los valores de muestras pares y los valores pares son actualizados con sumas ponderadas de las muestras impares. En el caso de la compresión. Sin pérdida - modo reversible- los resultados se redondean a valores enteros. Las fórmulas son:

$$y(2n+1) = x_{ext}(2n+1) - \left[\frac{x_{ext}(2n) + x_{ext}(2n+2)}{2} \right]$$

$$y(2n) = x_{ext}(2n) + \left[\frac{y(2n-1) + y(2n+1) + 2}{4} \right]$$

TESIS CON
FALLA DE ORIGEN

donde x_{ext} es la señal de entrada extendida y y es la señal de salida. El filtro 5/3 permite codificación y decodificación sin pérdida siempre y cuando los valores de la imagen descomprimida no son ajustados cuando caen fuera del rango dinámico completo. Para calcular las muestras impares se emplea una suma ponderada de las muestras pares (realmente se calcula un promedio) y las muestras pares, entonces, se calculan actualizando los valores de las muestras impares. Para mayor detalle ver la sección 2.1.6.

5. Cuantización

En este punto se determina el compromiso existente entre la tasa de compresión y la distorsión de la imagen, sin embargo, en la decodificación al descartar parte de los datos codificados, pueden variar la tasa como la calidad de la imagen.

La parte I del estándar utiliza una cuantización escalar uniforme con "dead-zone" en el origen, en la parte II se emplea cuantización trellis (TCQ).

En esta operación se pierde información a menos que el paso de cuantización sea 1 y los coeficientes sean enteros (caso del modo reversible).

Cuantización escalar:

$a_b(u, v)$ Coeficientes de la transformada de la sub-banda b

$q_b(u, v)$ Valor cuantizado de dicho coeficiente

$$q_b(u, v) = \text{sign}(a_b(u, v)) \left\lfloor \frac{|a_b(u, v)|}{\Delta_b} \right\rfloor$$

Δ_b paso de cuantización relativo al rango dinámico de la sub-banda b

Consideraciones:

- 5.1 Un tamaño de paso es permitido por sub-banda.
- 5.2 Aunque un paso de cuantización es permitido por sub-banda, el rango dinámico de la sub-banda depende del número de bits usados para representar la celda-componente de la imagen original y de la opción de transformación wavelet.
- 5.3 Todos los coeficientes cuantizados de la transformada serán signados, incluso cuando los componentes originales no tuvieran signo, ya que los coeficientes deben ser expresados en una representación signo-magnitud antes de la codificación.
- 5.4 En el modo reversible es forzoso que $\Delta_b=1$

6. Codificación por entropía.

Es necesario apuntar que la codificación de entropía del estándar JPEG2000 toma su base en el algoritmo EZW, presentado por Shapiro [S2], aunque tiene relación directa con el EBCOT [T2].

El modelo del contexto divide los bits de los coeficientes wavelet cuantizados en grupos con estadística similar. lo anterior con el objeto de que el codificador aritmético puede comprimirlos eficientemente. Se emplea un sistema de codificación aritmética que comprime los símbolos binarios relacionados a un modelo de probabilidad adaptivo asociado con cada uno de los 18 diferentes contextos de codificación definidos. El algoritmo de codificación es el MQ⁸ que proporciona compresión sin pérdida de cada uno

⁸ Basado en el utilizado en el estándar JBIG2 para imágenes binarias.

de los coeficientes wavelets seleccionados en cada recorrido, además, el mismo algoritmo administra la adaptación de los modelos de probabilidad condicional, véase la figura 3.1.

En esta etapa, los puntos importantes que deben ser resueltos son:

- ¿En qué orden serán codificados los bits, ya que es preferible codificar bits que aporten información no redundante?.
- ¿Cómo se representarán los bits seleccionados?.
- ¿Cómo se formará la cadena de bits transmitida?.

i) Orden del escaneo.

Una característica importante de la codificación en JPEG2000 es la naturaleza adaptiva de la secuencia en la cual los bits de diferentes muestras son codificados, la tendencia es codificar en primer lugar la información más valiosa en el sentido de la disminución de la distorsión de la imagen.

La codificación por entropía se realiza a nivel de los planos de bits de los bloques de código, iniciando por el plano correspondiente al MSB hasta el plano LSB. El recorrido de los planos inicia en la parte superior izquierda con los cuatro bits de la columna; a continuación los cuatro bits de la segunda columna hasta completar el ancho del bloque. Entonces los cuatro bits que siguen en la primera columna se examinan y se repite el proceso hasta recorrer todo el bloque.

ii) Contextos utilizados

Los modelos de probabilidad usados por el codificador corresponden a 18 diferentes contextos, agrupados en tres tipos de primitivas:

- 10 para la codificación de la significación,
- 5 para la codificación del signo y,
- 3 para la codificación del refinamiento.

La determinación del contexto toma en cuenta el estado de la significación y el estado de las 8 muestras vecinas. Cuando 4 muestras en una columna son candidatos para la codificación de significación en el plano p y, los 8 vecinos de cada una de las 4 muestras son en ese momento insignificantes, una primitiva de "Codificación Run Length" es usada y un solo bit es codificado para indicar si la columna está formada de ceros o no.

Las relaciones de causalidad de las que depende la determinación de los contextos del modelo son afectados por el orden adaptivo inducido por los recorridos en la codificación.

iii) Orden de codificación.

Cada bit del coeficiente es codificado en solamente una de los tres siguientes recorridos:

A) Primer recorrido.

Propagación del significado: Se intenta codificar las muestras que con mayor probabilidad lleguen a ser significativas, entonces, un bit es codificado si su locación es no significativa pero al menos uno de sus ocho vecinos es significativo. Cuando una de estas muestras llega a ser significativa, sus vecinos insignificantes se convierten en candidatos para el resto de los recorridos lo que induce a una onda de propagación de decisiones significativas cuando se presentan características importantes como bordes y regiones de textura.

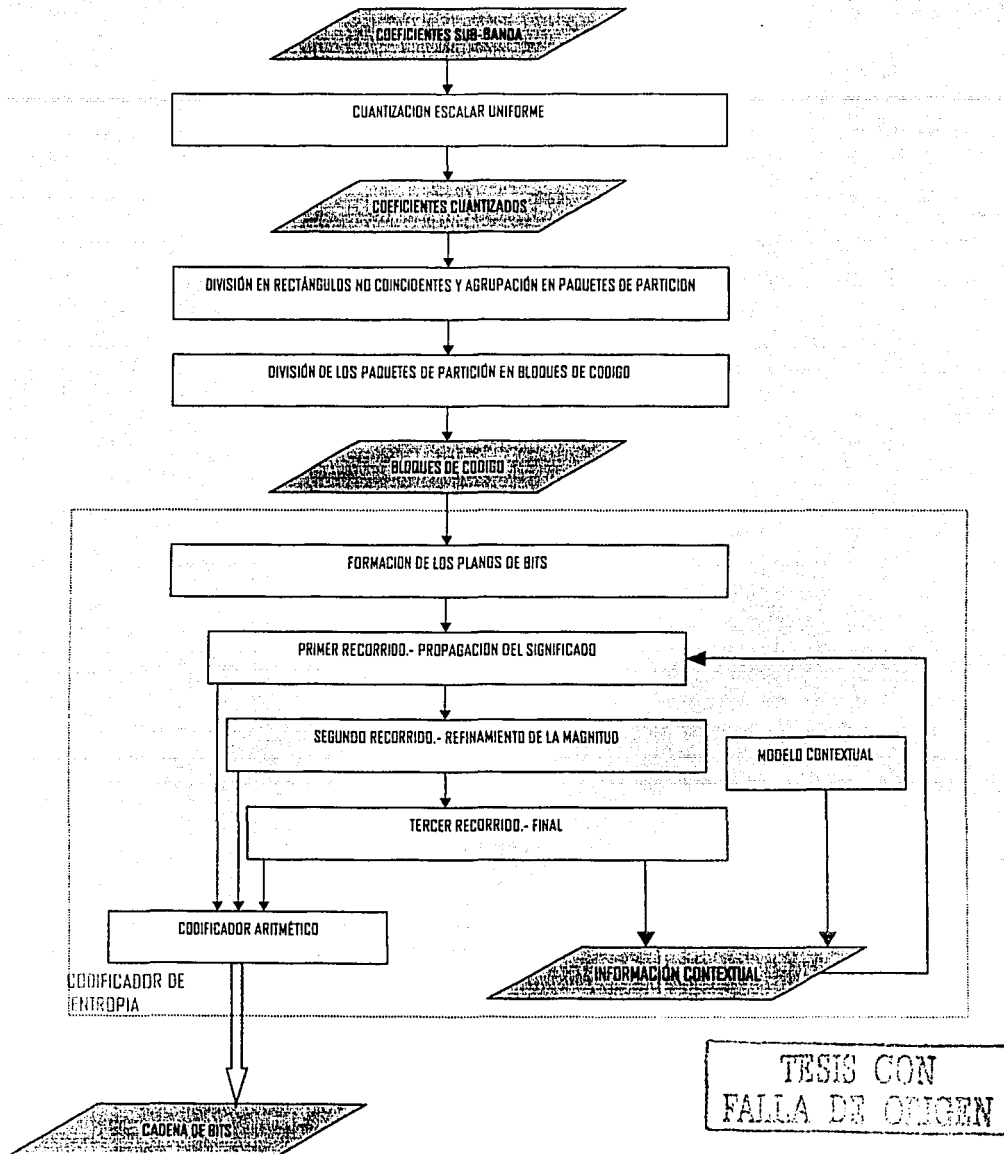


Figura 3.1. Proceso de cuantización y codificación por entropía.

Nueve contextos se crean con base en cuantos y cuales vecinos son significativos (Si un coeficiente es significativo entonces se le asigna a su posición en el contexto el valor de uno; de lo contrario se le asigna el valor cero). El mapeo del contexto también depende de la sub-banda a la que pertenece el bloque.

En otras palabras, este recorrido incluye solamente bits de coeficientes que fueron insignificantes hasta ese momento y tienen un contexto diferente de cero, el resto de los bits son saltados. El contexto es liberado al codificador aritmético junto con el bit analizado.

B) Segundo recorrido.

Refinamiento de la magnitud: Todos los bits que llegan a ser significativos en los planos anteriores son codificados, este recorrido incluye los bits de los coeficientes que ya han sido significativos (excepto aquellos que llegaron a ser significativos en el recorrido de propagación del significado).

El contexto que se usa es determinado por la suma del estado de significación de los vecinos horizontales, verticales y diagonales.

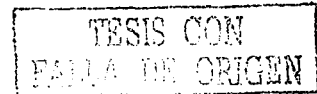
Estos son los estados que se dan a conocer al decodificador (no los estados usados antes del recorrido de decodificación de significación). Es mas, el estado es dependiente de si se trata del primer bit de refinamiento o no (el bit inmediato a los bits de significación y signo)

C) Tercer recorrido.

Final: En este recorrido se codifica todos los bits no codificados durante los recorridos anteriores (es decir, los coeficientes insignificantes con el valor del contexto igual a cero en el primer recorrido). Este recorrido no solo usa el contexto de la vecindad, sino un contexto de run-length, que se utiliza cuando las cuatro localidades de la columna del escaneo son insignificantes y cada localidad tiene vecinos insignificantes.

Las tablas de contexto pueden ser encontradas en [J1].

iv) Codificador aritmético.



En esta etapa la pareja decisión D y contexto CX (Etiqueta de contexto determinada durante los recorridos) es procesada para producir en la salida los datos comprimidos CD. Tanto D y CX son proporcionadas por la unidad de modelo en el codificador. CX selecciona el estimado de la probabilidad para ser usado durante la codificación de D.

La subdivisión recursiva del intervalo de probabilidad de la codificación de Elias es la base para el proceso de codificación aritmética binaria. Con cada decisión binaria, el intervalo de probabilidad actual es dividido en 2 subintervalos y la cadena de código es modificada (si es necesario) para que apunte a la base (el límite inferior) del subintervalo de probabilidad asignado al símbolo que ocurrió. Ya que el proceso de codificación involucra la adición de fracciones binarias en lugar de la concatenación de palabras de código enteras, las decisiones binarias más probables frecuentemente pueden ser codificadas a un costo menor al de un bit por decisión.

En cada partición en subintervalos, el subintervalo para el MPS es asignado arriba del subintervalo asignado al LPS. Por eso, cuando el MPS es codificado, el subintervalo LPS es añadido a la cadena de código. Esta convención en la codificación requiere que los símbolos sean reconocidos como MPS o LPS en lugar de 0 ó 1, en consecuencia, el tamaño del intervalo LPS y la detección (sentido) del MPS para cada decisión debe ser conocida para codificar esta decisión.

Ya que la cadena de código siempre apunta a la base del intervalo actual, el proceso de decodificación se encarga de determinar, para cada decisión, cual subintervalo es apuntado por los datos comprimidos. Esto es hecho también recursivamente, usando el mismo

proceso de subdivisión del intervalo. Cada vez que una decisión es codificada, el decodificador subtrae cualquier intervalo que el codificador haya añadido a la cadena de código. Por eso, la cadena de código en el decodificador es un apuntador dentro del intervalo actual con relación a la base de dicho intervalo.

Como se mencionó anteriormente, la característica del llenado de bits del codificador tiene la propiedad de que cualquier pareja de bytes liberados a la cadena de salida caen en el rango de 0 a FF8F, JPEG 2000 explota esta propiedad definiendo marcadores en el rango de FF90 a FFFF lo cual puede utilizarse para mejorar la protección contra errores y facilitar el análisis y la reorganización de la cadena de código. Además dado que $p_k \ll 1$, los eventos de renormalización ocurren raramente. Para la mayoría de los símbolos, el codificador MQ realiza solamente los siguientes pasos:

- 1.- $A \leftarrow A - p_k$

- 2.- $C \leftarrow C + p_k$

- 3.- Prueba de un bit para la condición de renormalización.

Las operaciones que consumen más recursos en cómputo ocurren solamente en la normalización, esto incluye:

- 1.- Transición de estados en el estimador de probabilidad adaptiva.

- 2.- El llenado de los bytes de código

- 3.- Los pasos de ajuste de A y C

La salida del codificador afecta el contexto para símbolos futuros. Así la corrida de símbolos consecutivos que pueden ser decodificados concurrentemente, de una manera práctica, es terminada con el primer bit decodificado que modifique el contexto para símbolos futuros, o por renormalización, lo que ocurra primero. Afortunadamente, las primitivas de codificación del JPEG 2000 tienen la propiedad de que los contextos para símbolos futuros son modificados solamente cuando una muestra insignificante llega a ser significativa, lo que varía el LPS, por lo que coincide con el evento de renormalización.

v) Formación de la cadena de bits

En este apartado se describirá la formación de la cadena de bits que finalmente se transmitirá o almacenará.

La generación de las cadenas de bits de los bloques de código son independientes entre si; en cada una de ellas, las fracciones de planos de bits (recorridos) son tratados como unidades indivisibles para la determinación de los puntos de truncamiento candidatos para la cadena, empleando la optimización tasa-distorsión**.

Durante el proceso de codificación, las longitudes y las distorsiones son calculadas y temporalmente almacenadas en la misma cadena de bits de los datos comprimidos.

Las cadenas de bits de cada bloque de código que forma un distrito (salida del codificador aritmético) comprenden el cuerpo de un paquete, cada uno de los cuales cuenta con su propio encabezado: una colección de paquetes (uno de cada distrito de cada nivel de resolución) comprende una capa; la cadena final es organizada como una sucesión de capas, en el estándar no se restringe la interpretación adoptada por el compresor en la reconstrucción de las capas de calidad.

Se añaden marcadores a la cadena de bits para permitir determinar la posición de los datos codificados con respecto a una localización espacial, de resolución o de calidad de la

** Aunque se soporta tamaños de paso de cuantización separadas para cada sub-banda, las propiedades tasa-distorsión de la representación embebida son insensibles a la selección de este paso.

imagen, esto con la finalidad de proporcionar: escalabilidad de calidad (SNR), escalabilidad espacial (resolución), recuperación a errores y codificación progresiva - con pérdida y sin pérdida -.

El principal encabezado de la cadena de salida describe las características de la imagen original y los estilos de descomposición y codificación utilizadas, además de la información necesaria para localizar, extraer, decodificar y reconstruir la imagen de salida con la resolución, fidelidad, ROI's u otras características deseadas.

Consideraciones:

- 6.1 El número restringido de contextos permite una adaptación rápida al modelo de probabilidad así como el decremento en el costo de los segmentos codificados independientemente.
- 6.2 En cada recorrido los contextos son creados y proporcionados al codificador aritmético
- 6.3 Los modelos de contexto son siempre reinicializados al inicio de cada bloque de código y el codificador aritmético es terminado al final de cada bloque. Esto es útil para la recuperación de errores.
- 6.4 Un modo de "lazy coding" es usado para reducir el número de símbolos que son codificados aritméticamente. En este modo, después del cuarto plano de bits codificado, el primero y segundo recorrido son incluidos como datos no comprimidos, es decir no se emplea el codificador MQ y solo es usado para el tercer recorrido de cada plano de bits.
- 6.5 Para la codificación de las regiones de interés los algoritmos que están considerados en el estándar son el método de escalamiento y el método Maxshift

3.6 Parte II del estándar

Etapa	Opciones disponibles
Cuantización	Además de la cuantización escalar con dead zone, se permite: Cuantización Trellis.- incrementa substancialmente el performance a varias tasas de bits, aunque se incrementa la complejidad del codificador.
Transformada wavelet	Además de la transformación wavelet en punto flotante (9.7) y la wavelet entera (3.5), se permite: Múltiples wavelets incluyendo las definidas por el usuario y cuyos coeficientes son especificados en la cadena de bits. También se permite descomposiciones diferentes a la de Mallat.
Transformación de los componentes	Se permiten transformadas arbitrarias y sus transformadas wavelets relacionadas.
Formato de archivo	Permite todos los espacios de color enumerados presentes en el estándar SPIFF, al igual que el conjunto completo de perfiles ICC para describir el espacio de color de una imagen. Además, la cadena de código puede ser dividida y distribuida en varias partes del archivo. Esto aumenta la facilidad con que imágenes de gran tamaño pueden ser editadas. Meta-datos también son definidos.

3.7 Notas sobre el estándar

Codificación por entropía

A pesar de que este algoritmo tiene su base en el algoritmo "zerotree" existen diferencias que mejoran la codificación en JPEg 2000:

- Cada bloque de código es codificado completamente independiente de los otros bloques de código en la misma u otra sub-banda, lo que permite que puedan ser codificados y decodificados en paralelo y en cualquier orden.
- Es posible el acceso espacial aleatorio al contenido de la imagen.
- Eficientes manipulaciones geométricas.
- Recuperación a errores.
- El usuario puede seleccionar arbitrariamente las contribuciones de cada bloque a las capas de calidad, mejorando el desempeño tasa-distorsión.

Condiciones para que el sistema sea reversible:

Los caminos que se pueden tomar en la codificación están definidos por la transformación a color y la transformada utilizada, sin embargo si la codificación reversible o sin pérdida es la deseada además de elegir correctamente las opciones en estos procesos se debe cumplir que el número de bits por muestra que presenta cada componente de la imagen de entrada debe ser idéntico al número de bits por muestra del componente correspondiente, de la imagen de salida o reconstruida.

3.8 Glosario de términos usados en el estándar.

Bloques de código.

Son los rectángulos no coincidentes en los que se divide un paquete de partición, su tamaño típico es de 64 x 64 y su tamaño mínimo 32 x 32 y son enumerados de acuerdo al paquete al que pertenecen.

Componente - celda

Unidad básica de la imagen original y reconstruida

Escalabilidad de resolución.

Habilidad para obtener imágenes de resolución más baja a partir de la misma cadena de bits.

Escalabilidad espacial.

Habilidad para obtener diferentes regiones de la imagen a partir de la misma cadena de bits.

LSB

Bit menos significativo.

MSB

Bit más significativo.

Paquete de partición (precincts)

Tres rectángulos espacialmente consistentes, un rectángulo de cada sub-banda en cada nivel de resolución.

ROI

Regiones de interés

Sistema reversible

Sistema sin pérdida de información.

SNR

Relación señal a ruido



Capítulo 4

SELECCIÓN DE WAVELETS

Este capítulo expone el desarrollo teórico de la justificación y la metodología de evaluación para la selección de wavelets. El objetivo de este capítulo es identificar las características de las funciones de escala y wavelet que mejoren el desempeño en la compresión de imágenes, y a partir de esto, determinar el conjunto de sistemas wavelets que serán consideradas en el estudio.

4.1 Objetivo de la selección de wavelets.

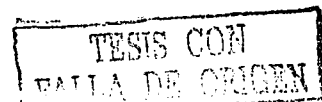
Encontrar una función wavelet que proporcione una descripción útil, eficiente y substancial de la señal de interés. Como el análisis del desempeño es dependiente de la aplicación en donde se emplee la wavelet, es necesario aclarar que ésta es la compresión con pérdida, sobre imágenes estáticas.

4.2 Metodología de evaluación.

Lo que se espera con la selección adecuada de la wavelet es poder mantener pequeña la distorsión debido a la cuantización, aunque se trabaje con ratios de compresión significativa. Para una clase de señales, el uso de una transformada unitaria particular - es decir, con una base wavelet determinada - puede ofrecer ventajas como la reducción del número de coeficientes cuya amplitud es mayor a un umbral fijo; este fenómeno recibe el nombre de compactación de energía y constituye una de las razones fundamentales del éxito de las técnicas de codificación por transformada. Otra ventaja se obtiene cuando los coeficientes son adecuados para la cuantización, ya que por medio de la codificación sub-banda se reducen los molestos efectos de ruido, producto de este proceso, que puede percibir el receptor; esto se logra principalmente porque al llevar a cabo el proceso de cuantización en cada banda, la señal reconstruida tiende a enmascarar esta clase de ruido.

La hipótesis que se sostiene a lo largo de este estudio es que el seleccionar funciones wavelets relacionadas a las ondas gaussianas, puede ser un factor determinante para alcanzar un alto desempeño en la compresión. Para probar lo anterior, la metodología será la siguiente:

1. Selección de los factores que se consideran relevantes en el desempeño de la transformada.
2. Selección de un conjunto de sistemas wavelets que de acuerdo a sus características aporten indicios sobre la importancia de la forma de onda de las funciones wavelet ψ y de escala ϕ .
3. Determinación de los valores de las variables en el estándar JPEG2000 para la inclusión de los sistemas wavelets seleccionados y realización de los experimentos correspondientes.
4. Evaluación objetiva y subjetiva de las imágenes comprimidas.
5. Conclusiones



Los primeros dos puntos son tratados en este capítulo, los puntos 3 y 4 son tema del capítulo 5 y las conclusiones son presentadas en el capítulo 6.

4.3 Selección del grupo de estudio

4.3.1 Variables consideradas.

Generalmente una aplicación basada en wavelets busca aproximar clases particulares de funciones con el menor número de coeficientes diferentes de cero. El diseño de la función wavelet debe entonces buscar maximizar el número de coeficientes $\langle f, \psi_{j,n} \rangle$ que son cercanos a cero; una función f tiene pocos coeficientes significativos, si la mayoría de los coeficientes wavelets de la menor escala (mayor resolución) son pequeños. Esto depende en mayor medida de la regularidad de f y de las siguientes propiedades de la wavelet:

- ✓ Ortogonalidad del sistema wavelet.
- ✓ Tamaño del soporte de las funciones wavelet.
- ✓ Número de momentos nulos* para las funciones wavelet.
- ✓ Regularidad de la función wavelet y de escala.
- ✓ Selectividad de frecuencia del sistema wavelet.
- ✓ Tamaño del filtros asociados al sistema wavelet.
- ✓ Fase de dichos filtros.
- ✓ Forma de onda de la wavelet.

En cuanto al número de niveles de descomposición y la estructura del árbol de descomposición, no se considerarán en el análisis porque la naturaleza de los algoritmos para la selección de la base de descomposición son independientes de la transformada que se utilice. Aunque es claro que su desempeño se vea afectado por las bondades de la transformación utilizada, estas estructuras o están fijas, como es el caso del algoritmo de Mallat, o bien, dependen de umbrales predefinidos como es el caso del algoritmo de la mejor base.

En el capítulo 2 se presentaron el origen y características de las principales familias de funciones que se emplean en las aplicaciones donde se descomponen las señales con bases wavelets. el capítulo 3 describe con algún detalle el estándar en el que las transformadas wavelets se emplearán e indica que en la extensión del JPEG 2000 (Parte II) se tiene la opción de indicar en la cadena de la imagen comprimida el conjunto de coeficientes de los filtros asociados a la transformación utilizada. Esto abre la posibilidad para la búsqueda de las wavelets más apropiadas para el proceso de compresión y, en determinado momento, para la imagen que se desea comprimir. Tomando en cuenta el contexto en el que se emplearán las wavelets, en la sección 4.3.3 se discutirán las familias que se incluirán en los experimentos.

4.3.2 Efecto de las variables en el desempeño de la wavelet.

Ortogonalidad:

Esta propiedad establece la relación directa entre la norma L_2 de una función y la norma de sus coeficientes wavelets. Para el caso en que se cumple la ortogonalidad:

$$\|f\| = \sqrt{\sum_{j,l} c_{j,l}^2}$$

* En inglés "Vanishing Moments".

En el caso biortogonal aunque esta igualdad no se cumple, ambas cantidades siguen siendo proporcionales.

Otra ventaja de los filtros ortogonales es que implementan una transformada unitaria entre la entrada y las sub-bandas. La característica útil de estas transformaciones es la conservación de la energía. En particular, la distorsión total es la suma de las distorsiones sub-banda y la tasa de bit total es la suma de todas las tasas de bit de sub-banda. En el caso de la no-ortogonalidad esto no se cumple y los algoritmos que fijan la tasa de bit para la transmisión o almacenamiento no se pueden aplicar directamente.

Soporte de las funciones asociadas al sistema wavelet:

La rápida convergencia de las funciones ψ y ϕ a cero, tanto en el dominio del tiempo como en el de la frecuencia, permite mejorar las propiedades de localización tiempo - frecuencia. Adicionalmente, partiendo de que si:

1. f tiene una singularidad aislada en t_0 y
2. t_0 está dentro del soporte de $\psi_{j,n}(t) = 2^{-j/2} \psi(2^{-j}t - n)$

entonces

$$\Rightarrow \langle f, \psi_{j,n} \rangle \text{ es probable que tenga una amplitud grande.}$$

Podemos concluir que si ψ tiene un soporte compacto de tamaño K , entonces en cada escala 2^j hay K wavelets $\psi_{j,n}$ cuyo soporte incluye t_0 a los que corresponderían K coeficientes de amplitud grande. Por esta razón, para minimizar el número de coeficientes de amplitud mayor se debe de reducir el tamaño del soporte de ψ . Minimizando, por ejemplo, el tamaño de la respuesta a impulso finita del filtro h , ya que si éste tiene un intervalo $[N_1, N_2]$ entonces ϕ también y ψ tiene un soporte de tamaño $N_2 - N_1$; de intervalo $[(N_1 - N_2 + 1)/2, (N_2 - N_1 + 1)/2]$.

En cuanto al esquema de codificación sub-banda, si la función de escala y wavelet son de soporte compacto, los filtros h y g son filtros FIR, lo que simplifica la implementación de la transformación.

Número de momentos nulos:

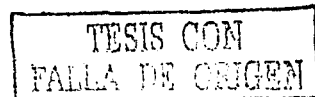
Decimos que ψ tiene p momentos nulos si:

$$\int_{-\infty}^{\infty} t^k \psi(t) dt = 0; \quad 0 \leq k < p \quad (4.1)$$

Esto significa que la función wavelet es ortogonal a cualquier polinomio de grado $p-1$. De hecho si f es regular y la función wavelet tiene suficientes momentos nulos, entonces los coeficientes wavelet son pequeños en escalas finas 2^j [M1]. Si f es localmente C^k , entonces, sobre un pequeño intervalo, f puede ser aproximado por un polinomio de Taylor de grado k . Si $k < p$, entonces los wavelets son ortogonales a este polinomio de Taylor lo que produce coeficientes de amplitud pequeña en las escalas más finas (tasa de convergencia de aproximaciones wavelet de funciones suaves).

Aunque el tamaño del soporte de una función y el número de momentos nulos son a priori independientes, las restricciones impuestas sobre las wavelets ortogonales implican que si ψ tiene p momentos nulos entonces el tamaño de su soporte debe de ser al menos de $2p-1$ [D1]. En el caso de las bases ortogonales, los siguientes enunciados son equivalentes:

- La función wavelet ψ tiene p momentos nulos.
- $\psi(\omega)$ y sus primeras $p-1$ derivadas son cero en $\omega = 0$.



- $h(\omega)$ y sus primeras $p-1$ derivadas son cero en $\omega = \pi$.
- Para cualquier $0 \leq k < p$,

$$q_k(t) = \sum_{n=-\infty}^{+\infty} n^k \phi(t-n) \text{ es un polinomio de grado } k$$

Regularidad de la función wavelet.

Un filtro ortogonal $h(\omega)$ con un cierto número de ceros en la frecuencia de aliasing ($\omega = \pi$ en el caso de dos canales) es llamado regular si su iteración tiende a formar funciones de escala y wavelet suaves, continuas o derivables. Por las ecuaciones fundamentales de escalamiento (Sección 2.1.4) es claro que las propiedades de regularidad de la wavelet se derivan directamente de las propiedades de la función de escala.

La importancia de esta propiedad en la codificación tiene dos partes fundamentales, primero, la presencia de muchos ceros en la frecuencia de aliasing puede mejorar la ganancia de codificación y segundo los artefactos de compresión podrían ser menos apreciables dado que tienden a ser más suaves las imágenes reconstruidas. La regularidad de ψ influye principalmente en el error introducido por la aplicación de umbrales o cuantización de los coeficientes wavelets. Cuando se reconstruye una señal a partir de sus coeficientes wavelets:

$$f = \sum_{j=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} f_j \psi_{j,n} \psi_{j,n}$$

un error ϵ añadido a un coeficiente $\langle f, \psi_{j,n} \rangle$ añadirá el componente wavelet $\epsilon \psi_{j,n}$ a la señal reconstruida. Si ψ es suave, entonces este componente es un error suave, esta característica es muy importante para la compresión ya que un error suave es menos visible en las imágenes reconstruidas que un error irregular, aunque presenten la misma energía.

Daubechies demostró que si ψ_k es ortonormal es imposible que ψ tenga caída exponencial y que al mismo tiempo sea C^∞ , con todas sus derivadas limitadas a menos que la función wavelet sea 0. Si el sistema empleado es biortogonal, la suavidad de función primaria (síntesis) es más importante para el error introducido que la función dual (análisis). En ambos casos una alto grado de suavidad corresponde a una mejor selectividad en frecuencia de los filtros como consecuencia directa del principio de incertidumbre

Finalmente la relación entre la regularidad y el número de momentos minimizados de ψ se puede enunciar como sigue: Si los filtros de escala son K -regulares, las siguientes sentencias son verdaderas y equivalentes:

- Los primeros K momentos de los filtros wavelets son cero.
- Los primeros K momentos de la función wavelet son cero.
- $h(\omega)$ tiene un cero de orden K en $\omega = \pi$.
- $\left. \frac{d^k}{d\omega^k} |h(\omega)|^2 \right|_{\omega=\pi} = 0$ para $k = 1, 2, \dots, 2K - 1$

Tamaño del filtro:

Anteriormente se mencionó el tamaño del soporte de las funciones, pero en la implementación de la transformada en un esquema de codificación sub-banda (caso del JPEG 2000) hay consideraciones adicionales que se debe indicar: Para que los filtros tengan buen rechazo fuera de banda (selectividad en frecuencia), o regularidad alta, se requiere filtros de gran orden; en contraparte, filtros con esta característica tienden a propagar los errores de codificación. Por ejemplo, los bordes pronunciados introducen distorsiones porque los canales de alta frecuencia son cuantizados de manera no muy fina. Si los filtros son grandes (y generalmente con respuesta impulso con varios cambios de signo), esto causa un artefacto desagradable conocido como "ringing" alrededor de los bordes .

Fase lineal de los filtros:

Si no se cumple esta condición la distorsión de fase alrededor de los bordes es muy visible. En un sistema FIR de dos bandas, la fase lineal y la ortogonalidad son mutuamente excluyentes; en la práctica se hace uso de sistemas separables de cuatro bandas. Con la ausencia de cuantización, la fase de los filtros no tiene sentido ya que el sistema tiene reconstrucción perfecta, esto mismo ocurre si se emplea una cuantización fina; en el caso de que la cuantización no sea fina – el paso de cuantización sea grande – la distorsión por fase es notoria. Para propósitos computacionales es preferible que los filtros tengan fase lineal debido a la simetría de los filtros.

Selectividad de frecuencia

Esa característica hace referencia a la capacidad de los filtros para discriminar las frecuencias y evitar la presencia de aliasing. Existe un compromiso con el tamaño del filtro ya que entre más corto es éste último, más se limita la selectividad de las frecuencias. Una ventaja de la selectividad de la frecuencia es que la ponderación perceptual de los errores es más fácil, ya que los errores son confinados a la banda donde ellos ocurren.

La selectividad para el filtro paso-bajas $H_0(e^{j\omega})$ se puede calcular con:

$$S_{H_0} = \int_0^{3\pi/8} (|H_0(e^{j\omega})| - |H_0(e^{j0})|)^2 d\omega + \int_{5\pi/8}^{\pi} |H_0(e^{j\omega})|^2 d\omega \quad (4.2)$$

Para el filtro paso-altas $H_1(e^{j\omega})$ el calculo se obtiene con:

$$S_{H_1} = \int_0^{3\pi/8} |H_1(e^{j\omega})|^2 d\omega + \int_{5\pi/8}^{\pi} (|H_1(e^{j\omega})| - |H_1(e^{j\pi})|)^2 d\omega \quad (4.3)$$

TESIS CON
FALLA DE ORIGEN

Forma de onda de la wavelet.

En realidad todas las características explicadas anteriormente afectan la forma de onda de la función wavelet y de la función de escala, es más, con la teoría de wavelets biortogonales no solo debemos de considerar una pareja de ondas wavelets sino de hecho, una pareja en el análisis y otra pareja para la síntesis.

La relación entre el grado de regularidad y el número de momentos nulos de la función wavelet está relacionada con las oscilaciones de la función. Las oscilaciones en el análisis causan que los productos escalares sean pequeños a escalas finas. Por otro lado la regularidad permiten la reconstrucción precisa de las señales.

Perceptualmente la forma de onda indica los eventos en la imagen a los que se le va a dar importancia, en una expansión lo ideal es que los coeficientes sean lo más significativos posibles y que con el menor número de ellos se concentre la mayor cantidad posible de información. La posibilidad de interpretar la información de estos coeficientes nos permite en gran medida la explotación eficiente de la información. En este aspecto, las funciones gaussianas y sus parientes cercanos han demostrado ser descriptores adecuados para los patrones relevantes en las imágenes.

Las representaciones de imágenes con filtros de gaussianas han mantenido un papel importante en las aplicaciones en visión. En el trabajo de Marr y Hildreth se emplean para la atenuación de ruido y la detección de bordes. Los operadores de laplaciana de gaussianas y diferencia de gaussianas han sido usadas para el análisis multi-resolución y se ha demostrado que las derivadas de gaussianas pueden representar de manera robusta las estructuras locales de las imágenes en múltiples escalas, tanto por sus propiedades en el dominio espacial como el de la frecuencia.

En la figura 4.1 se ilustra las interacciones entre los diferentes factores tratados anteriormente.

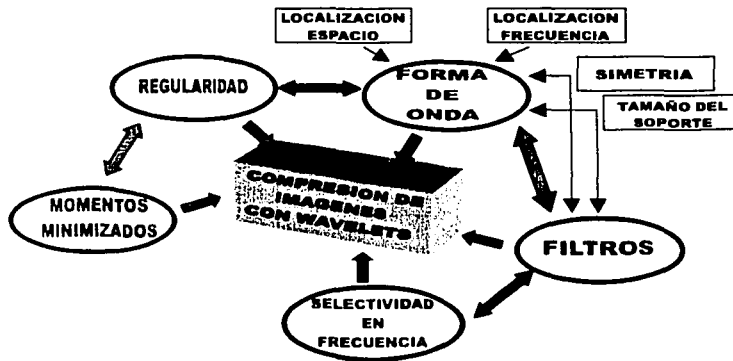


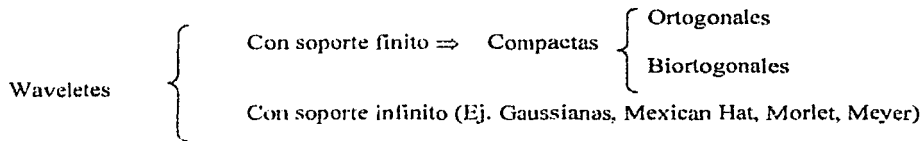
Fig 4.1 Factores que afectan el desempeño de un sistema wavelet

TESIS CON
FALLA DE ORIGEN

4.3.3 Análisis de las wavelets comparadas

La selección se limitará a los filtros wavelet separables como lo establece el estándar y considerará la literatura del estado del arte, en cuanto a que los sistemas utilizados sean populares y que siendo sometidos a otras técnicas de evaluación se haya encontrado que trabajan bien en la compresión de imágenes. A continuación se expondrán los argumentos teóricos de la selección efectuada, con el objeto de respaldar los resultados que se esperan. Una vez que se presenten las wavelets que se emplearan en las pruebas experimentales en el capítulo 5 se describirán las condiciones de los experimentos, los resultados obtenidos y cómo fueron evaluados. Finalmente en el capítulo 6 presentaré la discusión de los resultados.

Aunque no existe una clasificación única para las wavelets, de acuerdo a las características que se tratan en este trabajo, la siguiente es la más apropiada:



Del último conjunto de la clasificación, algunas no cuentan con una función de escala asociada, de hecho, de los ejemplos presentados en el cuadro la única con función de escala es la wavelet de Meyer (Sección 2.1.7).

La hipótesis que dio lugar a este trabajo gira en torno a que el uso de una base ortonormal con una wavelet madre suave y específicamente de una función relacionada con gaussianas o sus derivadas, tendrían mejor desempeño en la compresión de imágenes, no solo por la regularidad que presentan estas funciones y que las hacen útiles en la representación de imágenes (ya que la naturaleza de estas señales generalmente es suave a excepción de los bordes), sino porque al ser funciones simétricas o antisimétricas poseen una condición ventajosa para el manejo de los límites de la imagen ya que es posible extenderlos sin que el contenido de las imágenes se desplace entre sub-bandas.

Utilizando una extensión simétrica de la imagen, en lugar de una extensión periódica, se evitan introducir: (i) Discontinuidades producto de los bordes, que introducirían coeficientes de gran magnitud afectando directamente la cuantización, y (ii) artefactos de borde. Para que dentro del esquema de compresión explicado, se pueda tomar ventaja del anterior hecho, los filtros deben ser simétricos o antisimétricos, con lo que se garantizaría la no expansión de los coeficientes. Una consecuencia más relevante de que las wavelets no sean simétricas, en el procesamiento de imágenes, es que la visión humana es más tolerante a un error simétrico que a uno asimétrico. Desafortunadamente, la simetría perfecta sólo es posible para filtros wavelets complejos, wavelets biorortogonales, wavelets de soporte infinito y multi-wavelets.

Si bien debemos buscar que los filtros sean de tamaño pequeño, también se debe considerar que no pueden ser muy cortos porque implicaría tener menor suavidad de la función wavelet. Sin embargo, ¿qué es más conveniente seleccionar filtros de soporte pequeño o más suaves? De hecho la respuesta se encuentra en la discusión de las wavelets gaussianas, las cuales tienen el soporte infinito y la forma de onda adecuada.

Por lo anterior debería de ser seleccionada como primera opción, la wavelet gaussiana presentada en la sección 2.1.7, sin embargo, de esta selección surgen varios inconvenientes que a continuación trato con detenimiento.

Wavelets gaussianas

En realidad estas wavelet son frames wavelet porque no son bases ortonormales, otra característica de esta familia es que tiene soporte infinito. La falta de ortogonalidad imposibilitan su aplicación directa en la compresión de imágenes, ya que precisamente lo que queremos evitar es la redundancia en los coeficientes, además de la complejidad computacional derivada de la falta de ortogonalidad.

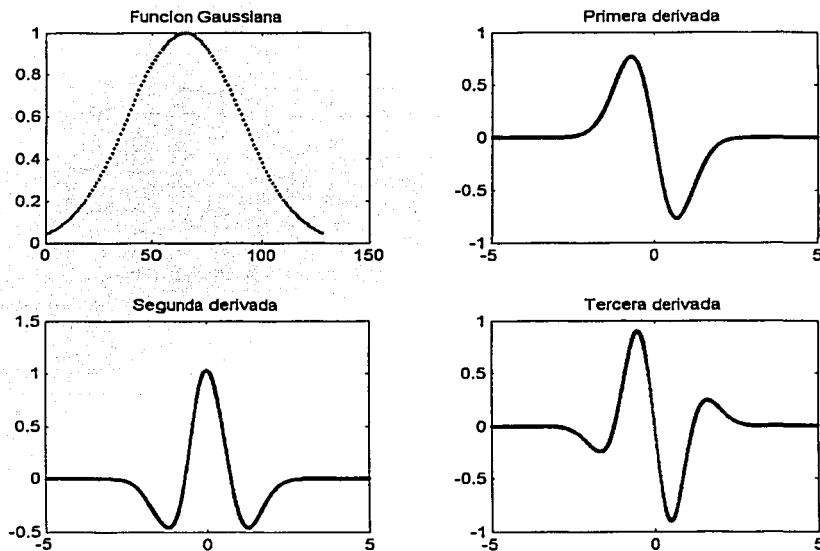


Fig. 4.2 Función gaussiana y las wavelets gaussianas de 1°, 2° y 3° orden.

Battle y Lemarié con base en el enfoque multi-resolución plantearon aproximar las señales con funciones B-spline, las cuales son muy parecidas a las función gaussiana aproximadas con coeficientes binomiales. A partir de esta función de escala, diseñaron una wavelet que al incorporar las condiciones de ortogonalidad que debían de cumplir, tanto la forma de onda de la función wavelet como de escala se veían afectadas; el resultado se conoce como la familia de wavelets ortogonales spline.

El truco de ortogonalización que se aplicó a las funciones spline se presenta en [D1] y las funciones obtenidas para el caso de la spline de grado 3 se presentan en la figura 4.3. Como se puede observar como consecuencia de la ortogonalización, las wavelets pierden su simetría y regularidad. Aunque las formas de la función de escala tienen parecido a la segunda y tercera derivada de una gaussiana (esto depende del grado de la función spline) existe un inconveniente aún más importante para el

uso de esta wavelet; su soporte, que originalmente era finito, para la versión ortonormalizada crece al infinito, dando lugar a una wavelet de soporte no compacto.

Aunque Lemarié probó que si la función wavelet tiene un soporte compacto entonces esta función necesariamente corresponde a una aproximación multi-resolución[M1], con este ejemplo podemos ver que lo contrario no se cumple, ya que un análisis multi-resolución no siempre se derivará en wavelets de soporte compacto o finito. Por otro lado, sí es posible construir wavelets que decaen como $|t|^{-1}$ al infinito y que no pueden ser asociadas a ninguna aproximación multi-resolución.

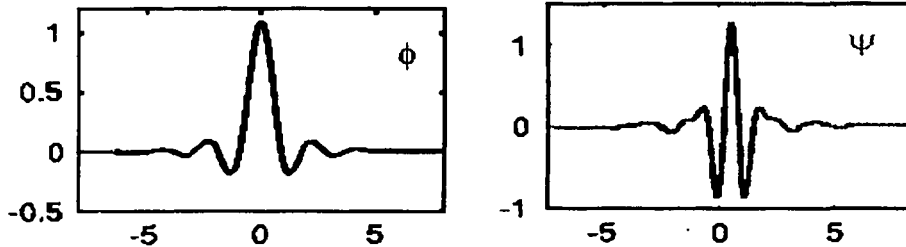


Fig. 4.3 Wavelet Battle-Lemarié cubic-spline

El hecho de que una wavelet este asociada a un esquema multi-resolución permite que el análisis de las funciones de interés pueda llevarse a cabo con un número finito de componentes y en este esquema es necesario introducir la función de escala. Sin embargo, la implementación de las transformaciones de soporte infinito implicarían el uso de filtros IIR. Entre las ventajas que obligan el uso de la codificación sub-banda en los wavelets utilizadas se encuentra que:

- Esta metodología resulta en una señal decimada al máximo, es decir, es posible reducir la tasa de muestreo, sin pérdida de información.
- Incluso si la salida de los filtros de análisis son cuantizados con pocos niveles (representar la muestra en 2 bits en lugar de 8 o 16 bits), la reconstrucción es aceptable, incluso con los errores de cuantización introducidos en la codificación [S1].
- Encontramos en la naturaleza muchos casos en los que se realiza un procesamiento similar. Por ejemplo, el oído y el ojo humanos, al menos en la primera etapa de la decodificación del sonido y la luz, realizan procesamientos con filtros muy relacionados al esquema codificación sub-banda.

Sin embargo, la condición de reconstrucción perfecta no es determinante para nuestra aplicación porque finalmente nos interesa la reconstrucción con pérdida, aunque la tendencia es tomar en cuenta la reversibilidad de la transformada, como lo puede constatar el número de estudios que se han hecho sobre transformadas wavelets definidas de enteros a enteros.[A3][A4][C2]

Siguiendo con el estudio de las wavelets gaussianas, la limitación real para su uso en el contexto del JPEG 2000 es que su soporte no es compacto por lo que no pueden ser relacionadas a bancos de filtros multi-resolución FIR. Las ventajas para el uso de las funciones de soporte compacto son:

- Tener una buena localización en el espacio.
- Poder ser asociados a algoritmos numéricos rápidos a través de la implementación con filtros FIR en un esquema sub-banda.

- Derivado del punto anterior, la ventaja de usar filtros FIR radica en; (i) el diseño de los filtros, ya que éste se reduce a resolver una ecuación que involucra que el cuadrado de la magnitud de la función de transferencia cumpla determinada especificación forzando a que el filtro sea de fase mínima. (ii) la propiedad de simetría de la respuesta a impulso de los filtros (propiedad que no se cumple con exactitud en el caso de los filtros IIR) equivalente a la fase lineal de los mismos

Sin embargo, los sistemas wavelets gaussianos son muy buena opción para el análisis de imágenes, eliminación del ruido, restauración, detección, reconocimiento y clasificación. [N1][R1][R2]. En cuanto a su aplicación en la compresión imágenes, en [U1] se presenta un algoritmo eficiente para el cálculo de la transformación utilizando wavelets de soporte infinito como la de Meyer y Battle-Lemarié, argumentando las buenas propiedades de localización espacio-escala y la ortogonalidad de las funciones base.

Ya que las propiedades que presentan las wavelets gaussianas imposibilitan su aplicación directa, la búsqueda se centra en seleccionar wavelets que compartan características con ella y que puedan ser efectivamente asociadas a esquemas de codificación sub-banda, tal como lo requiere el estándar. Las opciones dentro de la clasificación mostradas nos obligan a seleccionar funciones ψ de soporte finito ya sean ortogonales o biortogonales.

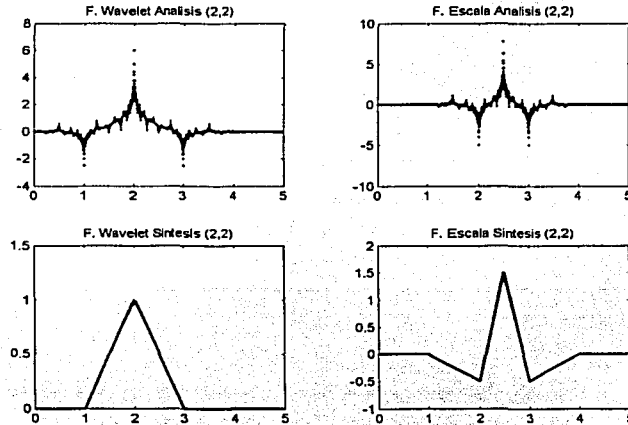


Fig. 4.4 Familia CDF B-Spline para $N=2$ y $\bar{N}=2$

TESIS CON
FALLA DE ORIGEN

Wavelets biortogonales.

Por los antecedentes que se presentaron de las wavelets, si se desea usar el mismo conjunto de filtros para el análisis y la síntesis y obtener una representación compacta de las señales, las wavelets deberán satisfacer las condiciones de ortogonalidad; mientras que si se desea utilizar dos filtros –asociados a dos conjuntos de wavelet distintos – uno para el análisis y otro para la síntesis, estos filtros deben cumplir con las condiciones de biortogonalidad. Las wavelets biortogonales se

basan en una pareja de funciones de escala $\phi(t)$ y $\tilde{\phi}(t)$ y dos funciones wavelet $\psi(t)$ y $\tilde{\psi}(t)$, lo que se traduce en flexibilidad para que los filtros de análisis y síntesis puedan diseñarse simétricos y por consiguiente, cumplir con un filtrado de fase lineal. La asociación entre las dos parejas de funciones se muestra con dos índices en las funciones, típicamente el primero representa el orden del filtro paso bajas para el análisis y el segundo el del filtro paso bajas de la síntesis. Para estos sistemas se cumple que el orden de los filtros puede ser par o impar y no necesariamente tienen que ser el mismo, la única restricción que se debe cumplir es que su diferencia debe ser par; o ambos son pares o ambos son impares. Los filtros considerados deben ser de fase lineal ya que no requieren compensación en fase y pueden utilizarse en estructuras de filtros piramidales.

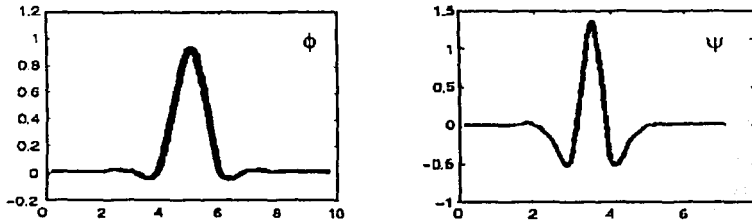
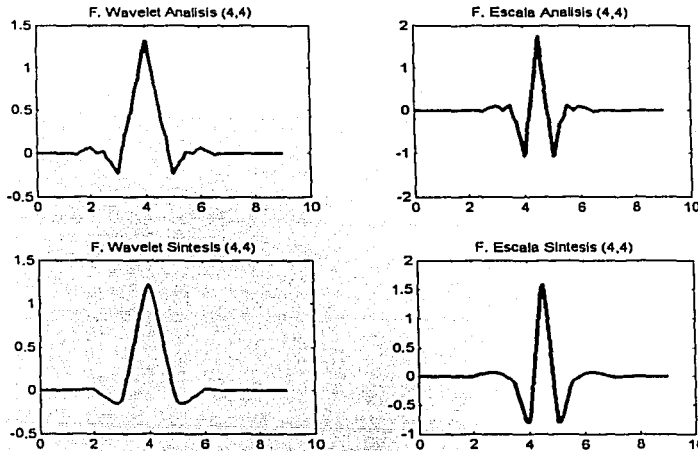


Fig. 4.5 Familia CDF B-Spline para $N=4$ y $\tilde{N}=2$

Para hacer una comparación de las diferencias entre preservar la forma de onda gaussiana sobre otras consideraciones seleccioné dos grupos de wavelets biortogonales; basadas en funciones B-spline (los filtros de análisis son la versión discreta de las funciones gaussianas, es decir, son funciones binomiales) y las wavelets biortogonales que conservan las características de las coiflets y que fueron diseñadas con base en el esquema lifting introducido en la sección 2.1.6.



WAVELETS BIORTOGONALES (4,4)

Fig. 4.6 Familia CDF B-Spline para $N=4$ y $\tilde{N}=4$

De la familia B-Spline CDF[†] serán incluidas aquéllas que sean más regulares en las funciones de síntesis ya que por [A2] sabemos que es más ventajoso conservar la regularidad en la etapa de síntesis comparada con la de análisis. En el caso de los sistemas biortogonales como los bancos de filtros FIR son de fase lineal, se presenta el caso en el que ya sea en la síntesis o en el análisis la función wavelet es regular, pero no ambas. Las gráficas de las funciones de síntesis escala y wavelet son presentadas en las figuras 4.4, 4.5 y 4.6

Para el resto de la selección de wavelets consideré las variables ya explicadas, cuando seleccionamos una wavelets nos enfrentamos al compromiso existente entre el número de momentos nulos y el tamaño de soporte de la wavelet. Si f tiene pocas singularidades aisladas y es muy regular entre singularidades, debemos seleccionar una wavelet con muchos momentos nulos para producir un gran número de coeficientes wavelets de magnitud pequeña. Si la densidad de las singularidades se incrementa, sería mejor disminuir el tamaño de su soporte aún a costa de reducir el número de momentos nulos, ya que las wavelets cuando se traslapan en las singularidades crean coeficientes de amplitud grande.

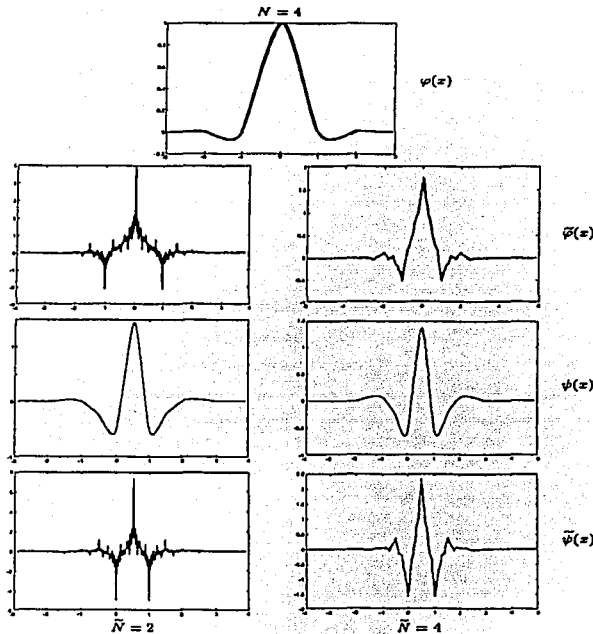


Fig. 4.7 Coiflets biortogonales $N=4$

Con respecto a las coiflets, refiriéndome a las wavelets ortogonales, las ventajas que aportan es que son menos asimétricas que la familia Daubechies y lo más importante es que un criterio de diseño es el número de momentos minimizados de la wavelet, lo que los hace similarmente efectivos para representar tanto los detalles como la aproximación de la imagen. Por los desarrollos de Daubechies en [D1] se sabe que para un ancho de soporte fijo de la función de escala y wavelet, o equivalentemente, para longitudes fijas de los filtros en el esquema sub-banda asociado, la selección de h_n que proporciona la regularidad máxima es diferente a la selección de la función con el máximo número N de momentos minimizados para la función wavelet; dado que los enfoques de diseño son distintos, parece interesante incluir una familia de coiflets en el estudio.

En el campo de la wavelets biortogonales, Sweldens presenta en [S5] una familia de coiflets biortogonales obtenidas a través del esquema lifting, tomando como base los filtros presentados por

[†] Ver sección 2.1.7

Deslauriers-Dubuc y la wavelet de Donoho. Los filtros Deslauriers-Dubuc tienen índices pares N y si son denotados por h^N sus propiedades se pueden expresar como:

- $h^N(\omega) + h^N(\omega + \pi) = 1$
- $h^N(-\omega) = h^N(\omega)$; es decir los filtros son simétricos.
- h^N tiene una raíz de orden N en π

La wavelet de Donoho tiene como función dual de escala la delta de Dirac en el origen, y la función wavelet:

$$\psi(x) = 2\phi(2x - 1) \quad (4.4)$$

La wavelet dual es una combinación lineal de deltas de Dirac:

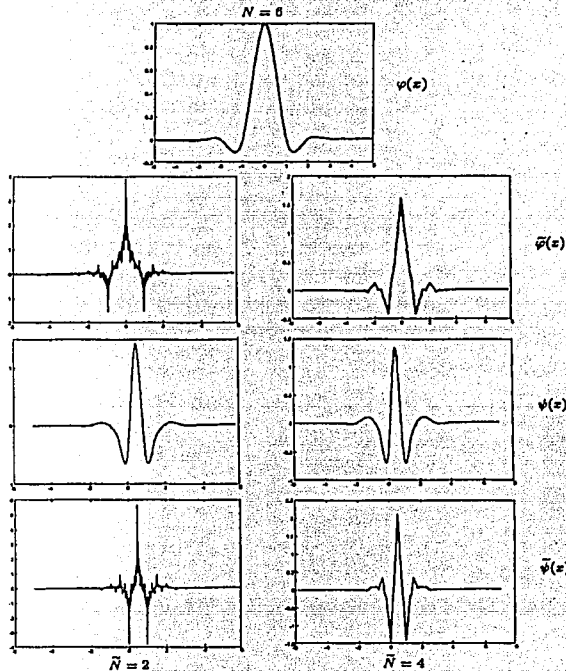
$$\tilde{\varphi}(\omega) = e^{-i\omega} \overline{h(\omega + \pi)} \quad (4.5)$$

donde h es un filtro que cumpla con:

$$h^N(\omega) + h^N(\omega + \pi) = 1 \quad (4.6)$$

Las desventajas de la wavelet de Donoho que se eliminan con el diseño de Sweldens son:

- Estas wavelets no tienen una integral cero por lo que no pueden formar una base Riesz para L_2 .
- $\tilde{h}(\omega)$, por ser una constante, introducía considerable aliasing.



Los miembros de esta familia que se consideraron corresponden a: (i) $N=4$ y $\tilde{N}=2$ (ii) $N=4$ y $\tilde{N}=4$ (iii) $N=6$ y $\tilde{N}=2$, los cuales presentan en las figuras 4.7 y 4.8.

Finalmente, se incluirán dos wavelets diseñadas en [A4], donde las formas de onda de sus funciones de síntesis son suaves pero los principales criterios de diseño son la simplicidad en la implementación y la reversibilidad de la transformación. Utilizando el esquema lifting, los requerimientos que cumplen estas wavelets son:

$$\left. \begin{aligned} |H_0(e^{j0})| &\neq 0 \\ |H_0(e^{j\pi})| &= 0 \\ |H_1(e^{j\pi})| &\neq 0 \\ |H_1(e^{j0})| &= 0 \end{aligned} \right\} \quad (4.7)$$

Fig. 4.8 Coiflets biortogonales $N=6$

donde $H_0(z)$ y $H_1(z)$ son los filtros de análisis paso-bajas y paso-altas, respectivamente. Además:

- El número de momentos minimizados para las funciones wavelet de análisis y síntesis deben ser mayor o igual a 2.
 - La selectividad en frecuencia en el análisis, tanto de $H_0(z)$ como de $H_1(z)$ debe ser menor o igual 0.25
 - Las longitudes de los filtros $H_0(z)$ y $H_1(z)$ son mínimas
- } (4.8)

Las funciones de las wavelets y de escala se muestran en las figuras 4.9 y 4.10.

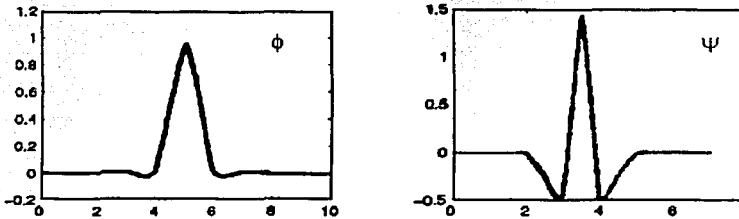


Fig 4.9 Wavelet reversible 5/11

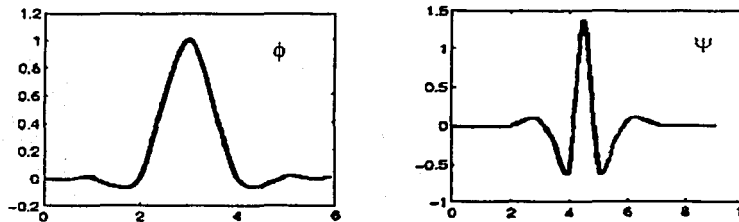


Fig 4.10 Wavelet reversible 13/7

Wavelets ortogonales

Como se ha hecho notar, la principal desventaja que presentan las wavelets cuando son ortogonales es que las funciones encontradas son extremadamente irregulares en el caso de la familia de Daubechies las funciones son incluso de naturaleza fractal[‡]. Sin embargo, las características por las que han sido una opción en las aplicaciones con wavelets son:

- Para un orden de la wavelet dado[§] (orden del filtro), Daubechies desarrolló las wavelets con regularidad máxima, es decir, las wavelets donde el número de momentos minimizados para $\psi(t)$ es máxima. (4.5)

[‡] Lo que es útil en el análisis aproximado o fractal.

[§] El orden de los filtros wavelets ortogonales siempre es par.

- Tiene buenas propiedades de compresión para los coeficientes wavelet pero no para los coeficientes de aproximación. Es decir, son muy buenas para representar de manera compacta los detalles de la señal, pero no para representar la aproximación de la señal en una resolución dada.
- Como ya se mencionó el número p de ceros en $h(\omega)|_{\omega=\pi}$ es igual al número de momentos nulos de ψ . Se puede observar que a priori esto no garantiza que al incrementar p mejoraremos la regularidad de la wavelet. Sin embargo para familias como las splines y las Daubechies, la regularidad de la wavelet se incrementa con el número de momentos nulos. Tchamitchian establece la relación entre el número de momentos minimizados de la función wavelet y la suavidad o diferenciabilidad de las funciones wavelet y de escala, indicando que entre mayor se el número de momentos minimizados, mayor será la regularidad de estas funciones. Es importante enfatizar que el número de momentos nulos y la regularidad de las wavelets ortogonales están relacionadas, pero es el número de momentos nulos y no la regularidad la que afecta la amplitud de los coeficientes wavelets en escalas finas, por lo que para efectos de compresión es mejor seleccionar el mayor número de momentos nulos.

Para la evaluación se incluirán las Daubechies de 1, 4 y 6 momentos minimizados (Figuras 4.10 y 4.11). La Daubechies con un momento minimizado corresponde a la wavelet Haar, que ya fue presentada en la sección 2.1.7. Esta es la wavelet más simple y la única wavelet ortogonal que tiene filtros de análisis y síntesis simétricos. Su uso es popular en situaciones con recursos computacionales limitados.

Para finalizar con la definición del conjunto de wavelets seleccionadas, la Tabla 4.1 presenta sus características relevantes.

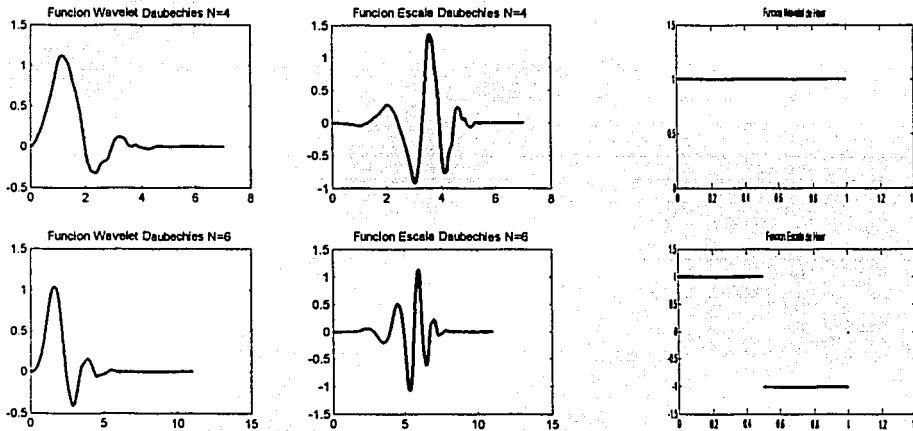


Fig 4.11 Familia Daubechies N=4 y N=6

Fig 4.12 Wavelet Haar

Nombre	Familia Wavelet	Ortogonal/ Biortogonal	Momentos nulos	Regularidad*	Selectividad frecuencia	Tamaño Filtro escala	Ecuaciones	Gráficas de funciones
HAAR	Daubechies	Ortogonal	1	0.0	Caida a razón de $1/\omega$	1	(2.10) (2.11)	Fig. 4.12
CDF(2,2)	B-Spline	Biortogonal	Análisis: 2 Síntesis: 2	Análisis:0.0 Síntesis:1.0	Paso bajas: 0.165 Paso altas: 0.192	Análisis:5 Síntesis:3	(2.27)	Fig. 4.4
CO(4,2)	Coiflets	Biortogonal	Análisis: 4 Síntesis: 2	Análisis:0.142 Síntesis: 2.0	Paso bajas: 0.129 Paso altas: 0.065	Análisis:9 Síntesis:7	(4.4), (4.5) (4.6)	Fig. 4.7
CO(6,2)	Coiflets	Biortogonal	Análisis: 6 Síntesis: 2	Análisis:1.245 Síntesis: 2.0	Paso bajas: 0.062 Paso altas: 0.025	Análisis:13 Síntesis: 11	(4.4), (4.5) (4.6)	Fig. 4.8
CO(4,4)	Coiflets	Biortogonal	Análisis: 4 Síntesis: 4	Análisis:0.841 Síntesis:2.0	Paso bajas:0.062 Paso altas:0.065	Análisis:13 Síntesis: 7	(4.4), (4.5) (4.6)	Fig. 4.7
DB4	Daubechies	Ortogonal	4	1.275	Caida a razón de $1/\omega^{1.275}$	7	(2.13)	Fig. 4.11
DB6	Daubechies	Ortogonal	6	1.888	Caida a razón de $1/\omega^{1.888}$	11	(2.13)	Fig. 4.11
CDF(4,4)	B-Spline	Biortogonal	Análisis:4 Síntesis:4	Análisis:1.039 Síntesis:1.701	Paso bajas:0.043 Paso altas:0.058	Análisis:9 Síntesis: 7	(2.27)	Fig. 4.6
5/11	Reversible	Biortogonal	Análisis:2 Síntesis: 2	Análisis:0.0 Síntesis:1.22	Paso bajas:0.165 Paso altas:0.131	Análisis:5 Síntesis: 11	(4.7) (4.8)	Fig. 4.9
CDF(4,2)	B-Spline	Biortogonal	Análisis: 4 Síntesis: 2	Análisis:0.0 Síntesis:2.142	Paso bajas:0.165 Paso altas:0.085	Análisis: 5 Síntesis: 11	(2.27)	Fig. 4.5
13/7	Reversible	Biortogonal	Análisis: 4 Síntesis: 2	Análisis:0.809 Síntesis:2.0	Paso bajas:0.029 Paso altas:0.065	Análisis:13 Síntesis: 7	(4.7) (4.8)	Fig. 4.10

Tabla 4.1 Cuadro de las wavelets seleccionadas.

* De la función wavelet.

† De la función de escala.

TESIS CON
 FALLA DE ORDEN

Capítulo 5

RESULTADOS

En el capítulo 4 se presentaron las transformaciones wavelet y los argumentos teóricos que hacen interesante su estudio. En este capítulo se presentan las condiciones de evaluación de este conjunto de transformaciones y los resultados obtenidos.

5.1 Software utilizado

El software utilizado es el modelo de verificación para el estándar JPEG 2000 parte I y II, versión 8.5, obtenido de <http://www.iv.optica.csic.es>. El lenguaje que se utilizó fue el C++, tomando todos los conceptos de la programación orientada a objetos. Conceptualmente, la transformada, el cuantizador y el codificador son objetos (implementados con estructuras que almacenan su estado y el flujo de datos entre las llamadas a funciones). Los autores llaman interfase al conjunto de funciones definidas por cada objeto (métodos) y es lo que permite la interacción con otros objetos o procesos.

Las etapas que sigue el proceso de compresión fueron implementadas en un modelo "push"; las muestras de una imagen son puestas en la etapa de análisis de wavelets una línea a la vez; la etapa de análisis pone las muestras sub-banda en la etapa de cuantización tan pronto como están disponibles (este proceso no espera a que toda la imagen sea transformada); de manera parecida la etapa de cuantización pone los índices del símbolo cuantizado en etapa de codificación; el cual por último pone los bits comprimidos en el objeto "cadena de salida" que se guarda en el archivo de salida para formar la imagen comprimida. El descompresor sigue el mismo proceso pero en orden inverso.

5.2 Variables que se fijaron en los experimentos

- Número de niveles de descomposición; la propiedad de aproximación sucesiva indica que conforme se añadan frecuencias más altas a la interpolación, se obtendrá una imagen de más alta resolución (esto equivale a generar más bandas o más señales de diferencia en la codificación subbanda y codificación piramidal respectivamente) Sin embargo el tamaño de las imágenes que se procesarán no es lo suficientemente grande para seleccionar un número grande de niveles de descomposición. En este caso preferí utilizar el valor pre-establecido en el estándar : 5 niveles.
- Estructura de descomposición; se utilizó el esquema de Mallat (Figura 2.7).
- Tasas de compresión; Se utilizó la cuantización TCQ por lo que el usuario puede establecer a qué tasa se llevará a cabo la compresión. El rango utilizado fue: 0.065, 0.125, 0.25, 0.50, 1 y 2 bits por píxel (bpp).
- Imágenes sobre las que se trabajó son la Bárbara de 512 x 512 píxeles y la Casa de 256 x 256 píxeles (Figura 5.1). Ambas son imágenes en niveles de gris y tienen originalmente 8 bpp

- ❑ Se evitó la división de la imagen en rejillas porque este proceso afecta directamente en los artefactos que presenta la imagen reconstruida, por lo que no se podrían discriminar de los artefactos que producen las transformadas, además dado el tamaño de las imágenes, no era primordial el uso de procesamiento en paralelo
- ❑ Los experimentos realizados utilizan el modo no reversible, el cual se explicó en el capítulo 3, sin embargo por condiciones del programa los filtros de análisis asociados a las wavelets se proporcionaron en su esquema lifting (sección 2.1.5). La Tabla 5.1 presenta la implementación de las funciones presentadas en la Tabla 4.1, de esta manera se definió la transformación en el software descrito en la sección anterior.



Fig 5.1(a) Bárbara



Fig 5.1(b) Casa

WAVELET	ESQUEMA LIFTING
HAAR	$d_i[n] = s_0[2n+1] - s_0[2n]$ $s_1[n] = s_0[2n] + 1/2 d_i[n]$
DB4*	$d_i^{(1)}[n] = s_0[2n+1] - \sqrt{3}s_0[2n]$ $s_1[n] = s_0[2n] - \sqrt{3}/4 d_i^{(1)}[n] + (\sqrt{3}-2)/4 d_i^{(1)}[n-1]$ $d_i[n] = d_i^{(1)}[n] + s_1[n+1]$
DB6*	$d_i^{(1)}[n] = s_0[2n+1] + 0.412286s_0[2n]$ $s_1^{(1)}[n] = s_0[2n] + 1.565136d_i^{(1)}[n-1] + 0.352387d_i^{(1)}[n]$ $d_i[n] = d_i^{(1)}[n] - 0.028459s_1^{(1)}[n] - 0.492151s_1^{(1)}[n+1]$ $s_1[n] = s_1^{(1)}[n] + 0.38962d_i[n]$
CO(4,2)	$d_i[n] = s_0[2n+1] - 9/16(s_0[2n] + s_0[2n+2]) + 1/16(s_0[2n-2] + s_0[2n+4])$ $s_1[n] = s_0[2n] + 1/4(d_i[n-1] + d_i[n])$
CO(4,4)	$d_i[n] = s_0[2n+1] - 9/16(s_0[2n] + s_0[2n+2]) + 1/16(s_0[2n-2] + s_0[2n+4])$ $s_1[n] = s_0[2n] + 9/32(d_i[n-1] + d_i[n]) - 1/32(d_i[n-2] + d_i[n+1])$
CO(6,2)	$d_i[n] = s_0[2n+1] - 75/128(s_0[2n] + s_0[2n+2]) + 25/256(s_0[2n-2] + s_0[2n+4]) - 3/256(s_0[2n-4] + s_0[2n+6])$ $s_1[n] = s_0[2n] + 1/4(d_i[n-1] + d_i[n])$
CDF(2,2)	$d_i[n] = s_0[2n+1] - 1/2(s_0[2n] + s_0[2n+2])$ $s_1[n] = s_0[2n] + 1/4(d_i[n-1] + d_i[n])$
CDF(4,2)	$d_i[n] = s_0[2n+1] + 1/16((s_0[2n+2] + s_0[2n-1]) - 9(s_0[2n+1] + s_0[2n]))$ $s_1[n] = s_0[2n] + 1/4(d_i[n] - d_i[n-1])$
CDF(4,4)*	$d_i^{(1)}[n] = s_0[2n+1] - 1.586(s_0[2n] + s_0[2n+2])$ $s_1^{(1)}[n] = s_0[2n] - 0.053(d_i^{(1)}[n] + d_i^{(1)}[n-1])$ $d_i[n] = d_i^{(1)}[n] + 0.883(s_1^{(1)}[n] + s_1^{(1)}[n+1])$ $s_1[n] = s_1^{(1)}[n] + 0.444(d_i[n] + d_i[n-1])$
5/11	$d_i^{(1)}[n] = s_0[2n+1] - 1/2(s_0[2n+1] + s_0[2n])$ $s_1[n] = s_0[2n] + 1/4(d_i^{(1)}[n] + d_i^{(1)}[n-1])$ $d_i[n] = d_i^{(1)}[n] + 1/32(s_1[n+2] - s_1[n+1] - s_1[n] + s_1[n-1])$
13/7	$d_i[n] = s_0[2n+1] + 1/16((s_0[2n+2] + s_0[2n-1]) - 9(s_0[2n+1] + s_0[2n]))$ $s_1[n] = s_0[2n] + 1/32(-d_i[n+1] - d_i[n-2]) + 9(d_i[n] - d_i[n-1])$

Tabla 5.1 Implementación de las wavelets seleccionadas.

$s_0[n]$: señal original

$d_i[n]$: coeficientes del filtro paso-altas después de la transformación wavelet

$s_1[n]$: coeficientes del filtro paso-bajas después de la transformación wavelet

$d_i^{(0)}[n]$, $s_i^{(0)}[n]$: Pasos de predicción y actualización intermedios

5.3 Resultados

Por cada imagen de entrada se obtuvieron 66 imágenes de salida. Los resultados para una tasa de compresión de 0.125 bpp se presentan en las figuras 5.2, 5.3.

Dado que el proceso de compresión con pérdida reduce la cantidad de datos almacenados o transmitidos, conduciendo a una pérdida real de información visual, es imperante conocer cuánta de la información de interés se perdió, identificando si es posible, la naturaleza de la misma. Los medios para lograr lo anterior se dividen en dos clases:

* Tanto el filtro como el esquema lifting tienen coeficientes irracionales, por lo que la aproximación dependerá de la precisión de la implementación.

- (1) Criterios de fidelidad objetivos: aquellos en los que la pérdida de información pueden ser expresada en función de la imagen original o de entrada y la imagen que después de ser comprimida se descomprime.
- (2) Criterios de fidelidad subjetivos: aquellos basados en la premisa de que los seres humanos son los usuarios finales de las imágenes comprimidas, por lo que es su evaluación la más apropiada para conocer la calidad de las mismas.

5.3.1 Análisis objetivo

Todas las codificaciones por transformada con pérdida introducen errores y distorsión. un punto interesante es encontrar para los diferentes métodos, clases y cantidad de artefactos que provocan. De la imagen comprimida se calculó la relación señal a ruido pico, la cual se define como [T1]:

$$PSNR = 10 \log_{10} \left(\frac{2^P - 1}{MSE} \right) \quad (5.1)$$

donde

$$MSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (x[i, j] - \hat{x}[i, j])^2$$

Las dimensiones de la imagen es $M \times N$ y esta representada por P bpp. La imagen original es denotada por x y la imagen comprimida por \hat{x} .

La tabla 5.2 muestra las PSNR obtenidas para cada una de las imágenes reconstruidas utilizando los diferentes sistemas wavelets. En la grafica 5.1 se presenta estas medidas de la diferencia con la imagen original Casa.raw para cada tasa de compresión, manteniendo fijo el resto de los parámetros. Mientras que en la grafica 5.2 se presentan los resultados para la imagen Bárbara.raw.

5.3.2 Análisis subjetivo

Aunque los errores pueden ser cuantificados, la magnitud de los mismos no siempre describirán la visibilidad de los errores que presentan, ya que el sistema visual humano no calcula ninguna de las métricas asociadas a la SNR, se hace necesario calificar la calidad de las imágenes a través de experimentos que involucren a usuarios finales que determinen la calidad visual, dependiendo de que tan bien reconozcan patrones familiares y entiendan el significado de los datos percibidos. Dado que el emitir un juicio sobre la calidad subjetiva de una imagen es un proceso cognoscitivo (depende de las experiencias individuales) donde las impresiones percibidas de los atributos básicos de la imagen juegan un papel importante, también se investigó cuáles eran los atributos que los observadores más ponderaban en su evaluación; siendo que éstas no se obtuvieron individualmente para cada factor, sino como una métrica de la calidad global de la imagen.

El método generalizado consiste en mostrar una imagen típica descomprimida a un conjunto de observadores y promediar sus observaciones: en estos casos las evaluaciones son hechas usando un rango absoluto de escalas. [G3] Otro método consiste en comparar dos imágenes, presentándolas simultáneamente, entonces se le pide al observador que emita una evaluación que consiste en un número positivo o negativo; tomando una referencia el observador determina cual de las dos imágenes es mejor y en que cantidad, por ejemplo, si la imagen a la izquierda es mejor la escala será negativa en caso contrario será positiva. [D3]

En este trabajo, los experimentos se realizaron reuniendo a un grupo de 5 observadores no expertos; del conjunto de imágenes comprimidas con una tasa de bpp menor a 2, se seleccionaron aleatoriamente 80 imágenes, al observador se le presentaron 100 imágenes con repeticiones; finalmente se les solicitó que asignaran a cada imagen una calificación entre 1 y 10. La validez de esta escala ha sido estudiada en [R3] y [R4].

Finalmente se preguntó a los sujetos los parámetros que tomaron en cuenta al momento de realizar sus evaluaciones, coincidiendo en que fundamentalmente era la nitidez de la imagen el factor con mayor ponderación; su atención la centraron en la preservación de los detalles como son los bordes y las texturas.

Con las calificaciones obtenidas se obtuvieron promedios para cada tipo de wavelet y tasa de compresión. Adicionalmente se obtuvieron la calificación mínima y máxima para cada tasa de compresión, y el promedio por sistema wavelet para indicar el desempeño global del sistema. Los resultados se muestran en la tabla 5.3. Las gráficas 5.3 y 5.4 muestran el desempeño subjetivo de las wavelets, mostrando la desviación estándar de cada promedio y finalmente en las gráficas 5.5 y 5.6 se presentan las comparaciones de estas métricas para la Casa y Bárbara, respectivamente.

5.3.3 Comparación entre la calidad subjetiva y la calidad objetiva.

Como último paso se elaboraron diagramas de dispersión entre los dos criterios considerados en la comparación, se consideraron cada una de las imágenes procesadas y los resultados obtenidos con ambas imágenes. Con esto se intenta medir la fuerza de la relación existente entre la métrica objetiva (PSNR) y la métrica subjetiva (considerada como el promedio de las evaluaciones emitidas). Los diagramas de dispersión se presentan en las gráficas 5.7, 5.8 y 5.9.

Adicionalmente se calculó los coeficientes de correlación, asumiendo una relación lineal entre las dos variables (gráfica 5.7); para la imagen Casa.raw $\rho = 0.93$, mientras que para Bárbara.raw $\rho = 0.98$, donde:

$$\rho_{xy} = \frac{\text{cov}(x, y)}{\sigma_x \cdot \sigma_y}; -1 \leq \rho_{xy} \leq 1 \quad (5.2)$$

$$\text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y)$$

Se debe aclarar que las imágenes comprimidas a 2 bpp no fueron consideradas en el análisis subjetivo por no presentar diferencias significativas. El análisis detallado de los resultados se presenta en el capítulo siguiente.

TABLA 5.2 : RESULTADOS CUANTITATIVOS

		PSNR EN dB PARA CASA.RAW (a)									
WAVELET	HAAR	CDF 2.2	CO 4,2	CO 6,2	CO 4,4	DB4	DB6	CDF 4,4	5/11	CDF 4,2	13/7
TASA[bpp]											
2	43.0432	45.2468	45.2424	45.1801	45.4447	42.1099	37.2298	45.3111	45.2614	45.1768	45.4275
1	35.1361	37.4511	39.3465	39.1636	37.8414	37.0408	29.29	37.9546	39.587	39.3008	37.8919
0.5	32.5862	34.2549	34.5065	34.3677	34.6801	31.9747	27.0759	34.7032	34.5178	34.412	34.6909
0.25	29.7215	31.4692	31.1838	31.3722	31.5889	29.856	24.7055	31.7666	31.4744	31.3766	31.5641
0.125	26.2096	28.3068	28.2952	28.2479	28.4452	25.4449	20.9359	28.674	28.2489	28.2065	28.3611
0.0625	22.6035	23.287	22.2238	22.302	23.5297	21.2059	18.9245	23.5787	22.1537	22.2016	22.3652
PROMEDIO	31.59335	33.3528167	33.4665667	33.4389167	33.5883333	31.1055333	26.3602667	33.6651833	33.5405333	33.44575	33.3834667

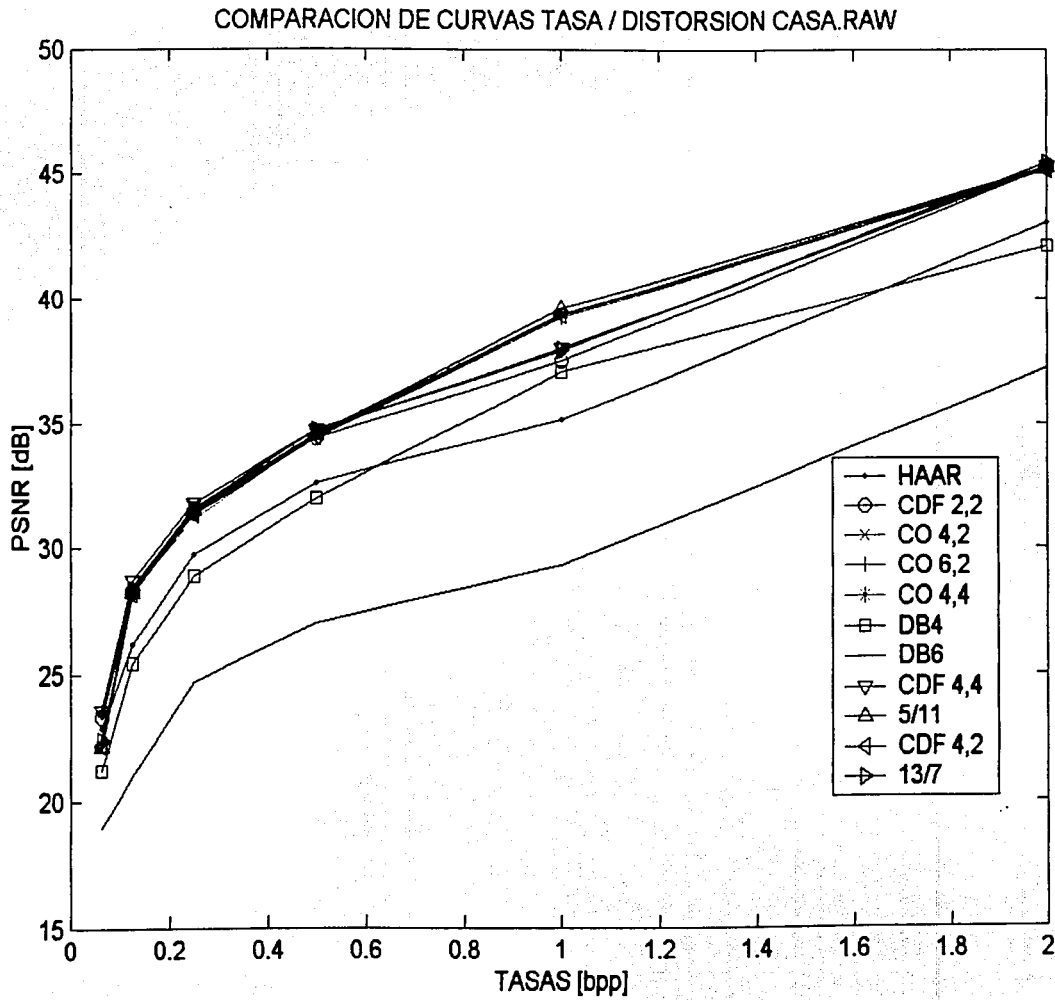
		PSNR EN dB PARA CASA.RAW (b)			
TASA[bpp]		MINIMO		MAXIMO	
2		DB6	37.2298	CO 4,4	45.4447
1		DB6	29.29	5/11	39.587
0.5		DB6	27.0759	CDF 4,4	34.7032
0.25		DB6	24.7055	CDF 4,4	31.7666
0.125		DB6	20.9359	CDF 4,4	28.674
0.0625		DB6	18.9245	CDF 4,4	23.5787
PROMEDIO		DB6	26.3602667	CDF 4,4	33.6651833

		PSNR EN dB PARA BARBARA.RAW (c)									
WAVELET	HAAR	CDF 2.2	CO 4,2	CO 6,2	CO 4,4	DB4	DB6	CDF 4,4	5/11	CDF 4,2	13/7
TASA[bpp]											
2	36.7042	38.8942	40.6428	41.3094	40.9326	34.9644	36.6973	40.5744	39.8334	40.6582	40.8208
1	33.7328	35.5357	35.5661	32.9092	36.2543	32.8043	31.7352	36.9024	35.0844	35.9605	35.5594
0.5	28.9028	31.1694	31.5371	31.9956	32.2142	27.8235	28.2675	32.3175	31.2657	31.4627	32.3614
0.25	25.938	27.4602	27.6531	27.7064	28.1069	24.9818	25.4254	25.3164	27.5932	27.5732	28.1942
0.125	23.9468	24.6795	24.8215	24.5523	24.5174	22.7119	23.2451	25.214	24.6364	24.5411	25.0779
0.0625	22.547	23.1079	23.1432	23.1608	23.2273	22.1195	21.8451	23.462	23.1212	23.1215	23.246
PROMEDIO	28.6285333	30.1411167	30.5941333	30.27125	30.9412833	27.5575667	27.8694333	31.1311167	30.4228833	30.5528667	30.8766167

		PSNR EN dB PARA BARBARA.RAW (d)			
TASA[bpp]		MINIMO		MAXIMO	
2		DB4	34.9644	CO 6,2	41.3094
1		DB6	31.7362	CDF 4,4	36.9024
0.5		DB4	27.8235	13/7	32.3614
0.25		DB4	24.9818	CDF 4,4	28.3164
0.125		DB4	22.7119	CDF 4,4	25.214
0.0625		DB6	21.8451	CDF 4,4	23.462
PROMEDIO		DB4	27.5675667	CDF 4,4	31.1311167

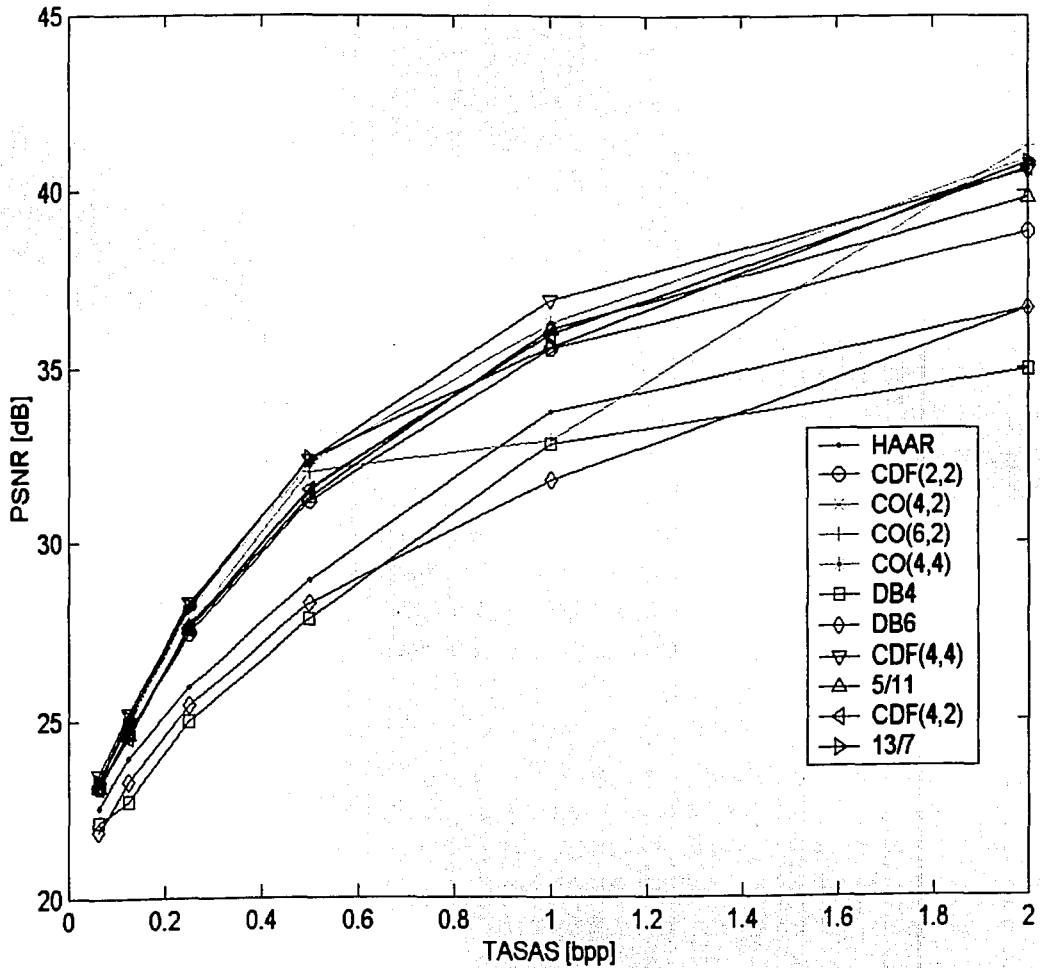
TESIS CON
 FALLA DE ORIGEN

TESIS CON
FALLA DE ORIGEN



Grafica 5.1 Comparación de calidad objetiva para la casa

COMPARACION DE CURVAS TASA / DISTORSION PARA : BARBARA512.RAW



Grafica 5.2 Comparación de calidad objetiva para barbara

TESIS CON
 FALLA DE SERVICIO

TABLA 5.3 : RESULTADOS CUALITATIVOS

CALIFICACION PARA CASA.RAW (a)

WAVELET TASA[ppp]	HAAR	CDF 2,2	CO 4,2	CO 6,2	CO 4,4	DB4	DB6	CDF 4,4	5/11	CDF 4,2	13/7	PROMEDIO
1	7.92	7.75	8.50	8.50	8.83	8.33	6.25	8.67	8.58	8.75	8.45	8.41
0.5	7.00	7.38	7.23	7.30	7.32	6.78	6.97	7.50	7.42	6.72	7.33	7.17
0.25	4.67	4.95	4.92	5.20	5.88	4.42	5.07	4.63	5.63	5.13	5.23	5.07
0.125	2.67	3.00	3.35	3.53	3.45	1.67	2.32	3.47	3.63	3.25	3.35	3.06
0.0625	1.37	1.45	1.26	1.28	1.72	1.03	1.45	1.92	1.78	1.59	1.51	1.49
PROMEDIO	4.72	4.91	5.05	5.16	5.44	4.45	4.80	5.24	5.41	5.09	5.18	5.04

CALIFICACION PARA CASA.RAW (b)

TASA[ppp]	MINIMO	MAXIMO
1	CDF 2,2	7.75 CO 4,4
0.5	CDF 4,2	6.72 CO 4,4
0.25	DB4	4.42 CO 4,4
0.125	DB4	1.67 5/11
0.0625	DB4	1.03 CDF 4,4
PROMEDIO	DB4	4.45 CO 4,4

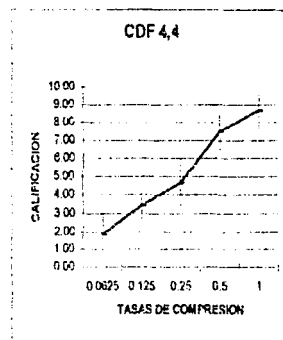
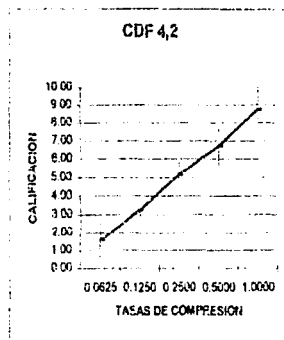
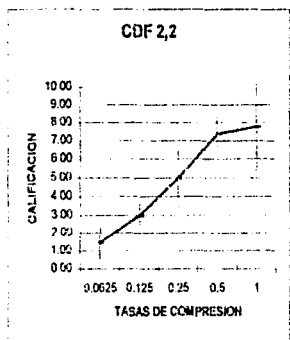
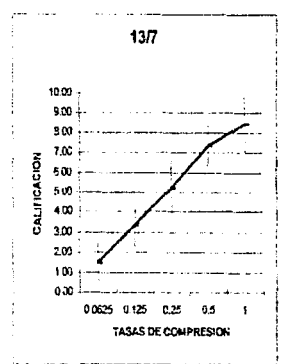
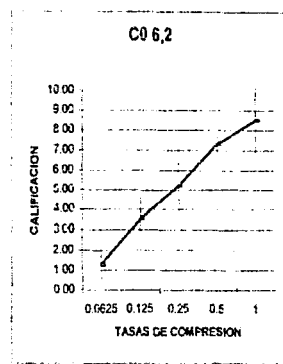
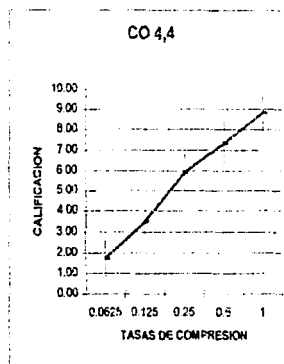
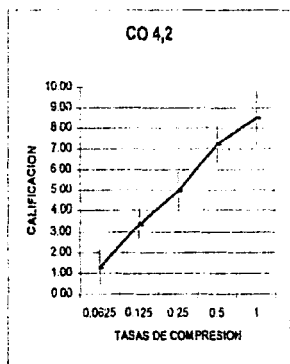
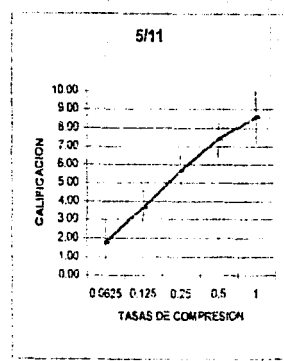
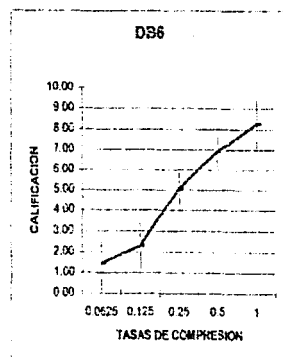
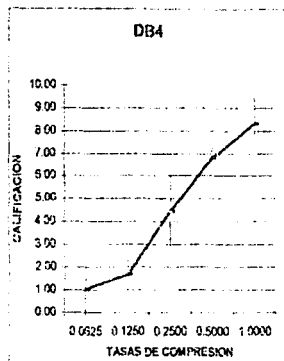
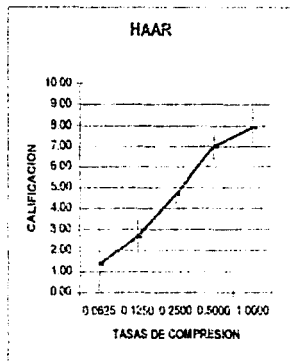
CALIFICACION PARA BARBARA.RAW (c)

WAVELET TASA[ppp]	HAAR	CDF 2,2	CO 4,2	CO 6,2	CO 4,4	DB4	DB6	CDF 4,4	5/11	CDF 4,2	13/7	PROMEDIO
1	8.08	8.58	8.53	8.49	8.48	8.17	8.33	9.07	8.92	8.93	8.82	8.64
0.5	6.58	7.53	6.83	7.27	7.27	6.10	7.28	7.45	7.48	7.35	7.52	7.16
0.25	5.75	6.22	5.98	6.32	6.46	5.13	6.47	6.47	6.17	6.16	6.32	6.18
0.125	4.38	5.08	4.98	5.07	5.30	3.30	5.20	5.06	5.33	4.43	4.93	4.82
0.0625	3.92	4.55	4.63	4.67	4.37	3.00	4.55	4.57	5.38	4.73	4.42	4.44
PROMEDIO	5.74	6.53	6.21	6.36	6.36	5.14	6.47	6.51	6.66	6.33	6.42	6.25

CALIFICACION PARA BARBARA.RAW (d)

TASA[ppp]	MINIMO	MAXIMO
1	HAAR	8.08 CDF 4,4
0.5	DB4	6.10 13/7
0.25	DB4	5.13 CDF 4,4
0.125	DB4	3.36 5/11
0.0625	DB4	3.00 5/11
PROMEDIO	DB4	5.14 5/11

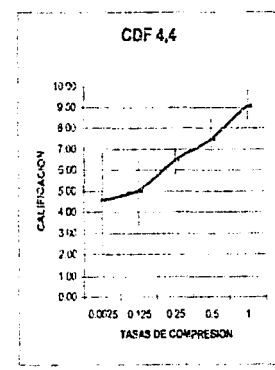
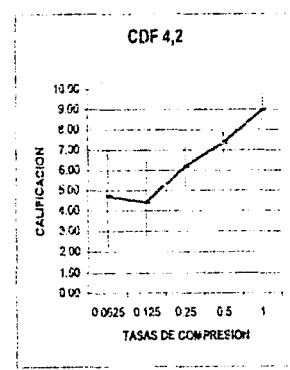
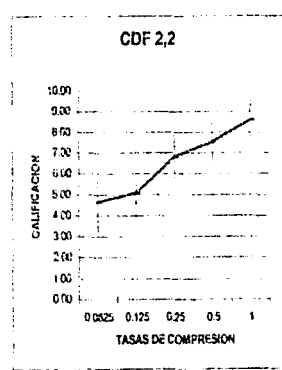
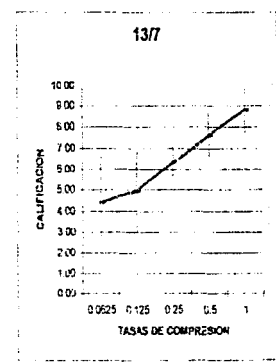
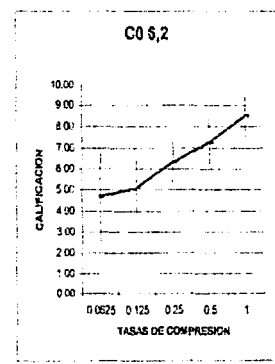
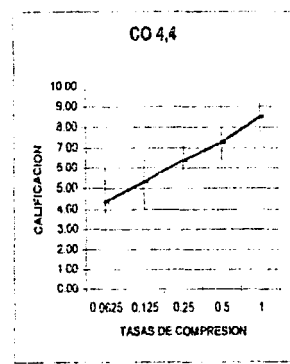
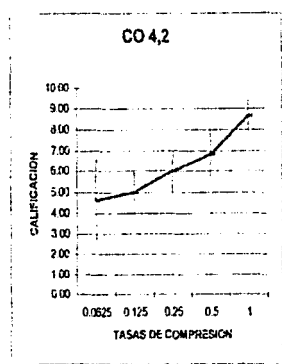
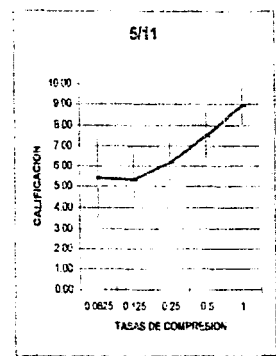
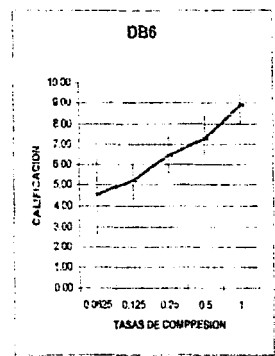
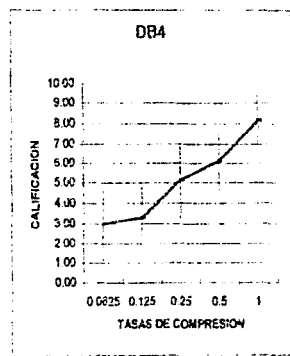
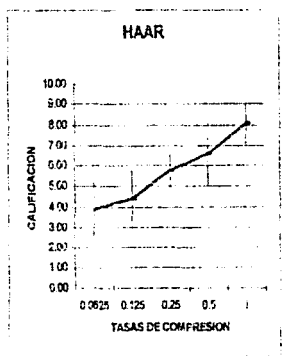
TESIS CON
 FALTA DE ORIGEN



Grafica 5.3 Calidad subjetiva para la casa: valores medios y desviación estándar

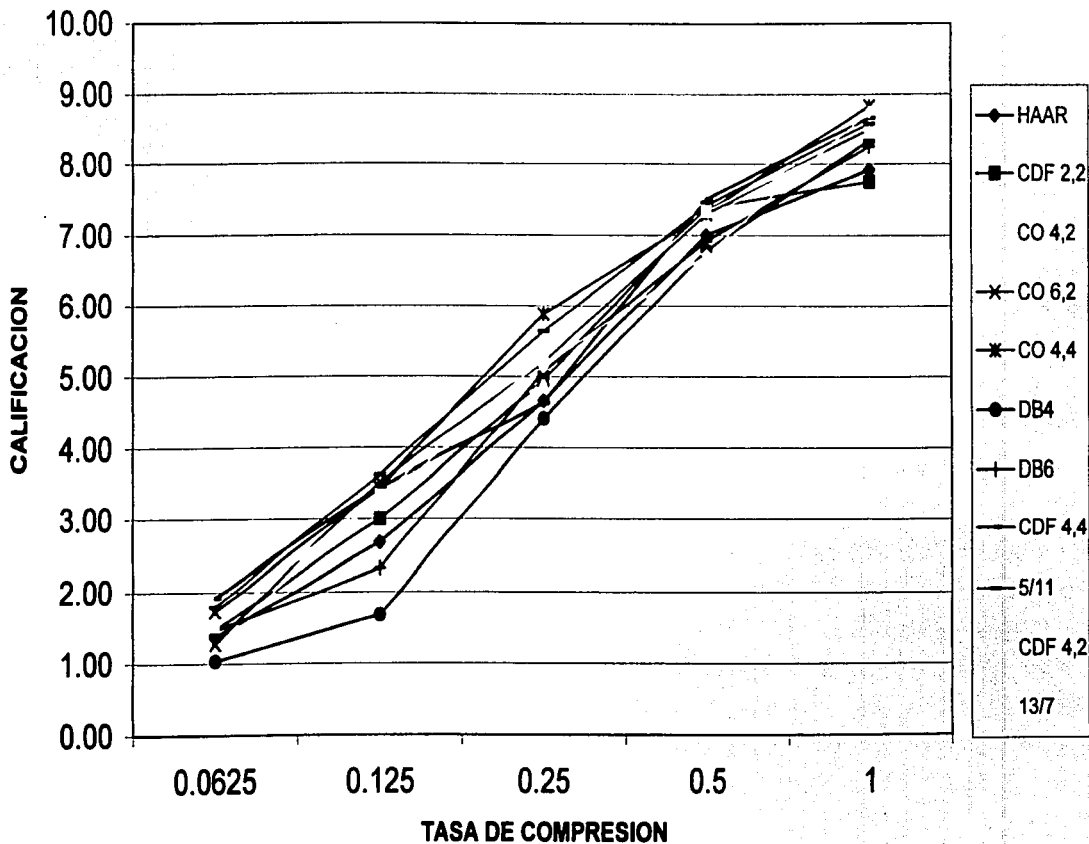
TESIS CON
 CALIFICACION
 NO CUMPLEN

TESIS CON
 FALLA DE ORIGEN
 MEDIO DE VITRE



Grafica 5.4 Calidad subjetiva para barbara: valores medios y desviación estandar.

CALIDAD SUBJETIVA: CASA.RAW

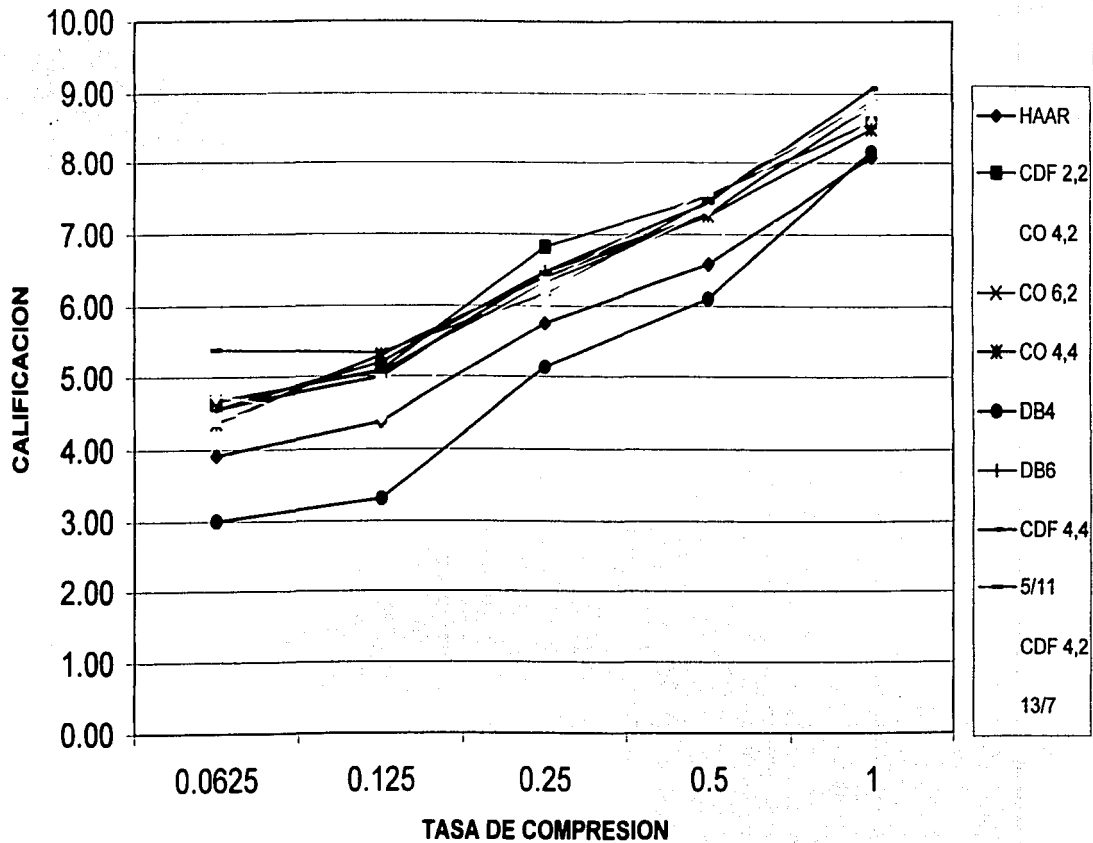


Grafica 5.5 Comparación de la calidad subjetiva de todos los sistemas wavelets para casa.raw

TESIS CON
 FALLA DE ORIGEN

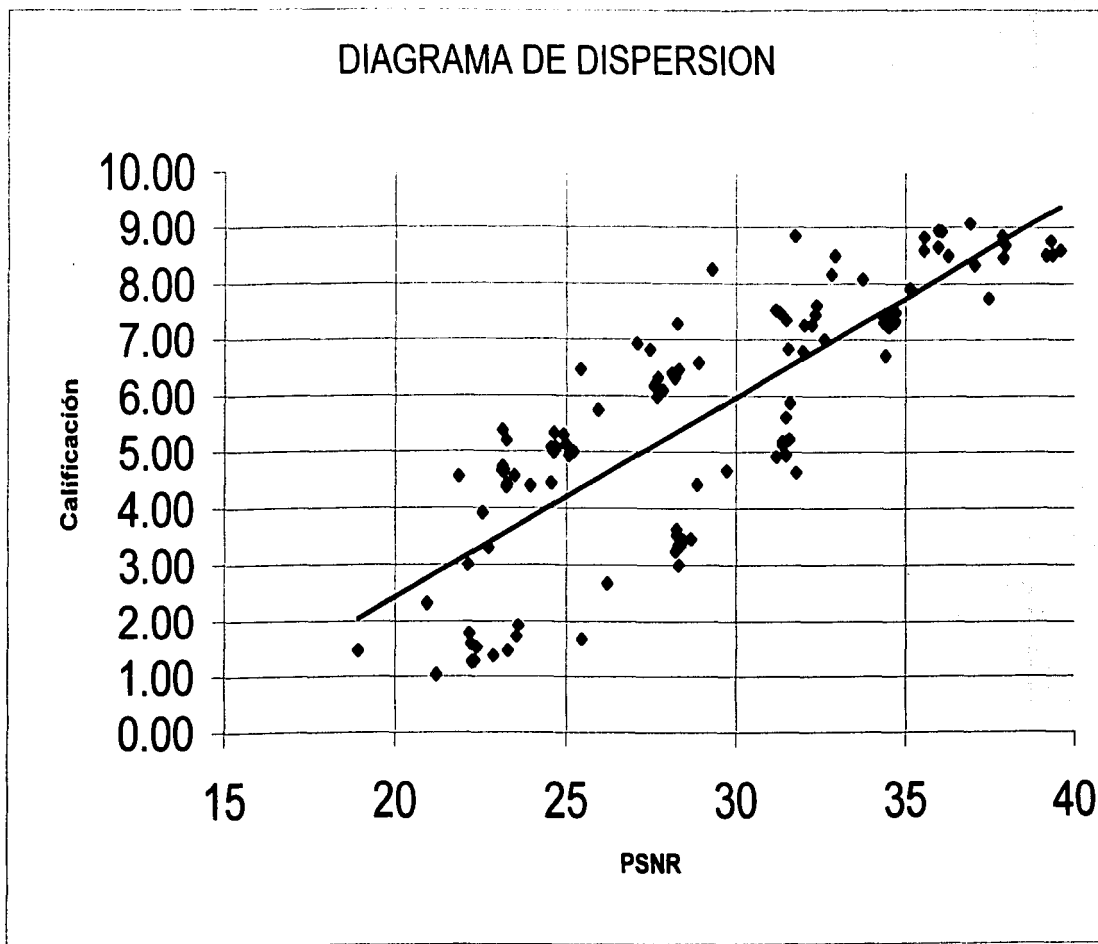
TESIS COM
FALLA DE ORIGEN

CALIDAD SUBJETIVA: BARBARA.RAW



Grafica 5.6 Comparación de la calidad subjetiva de todos los sistemas wavelets para barbara.raw

TESIS CON
FALLA DE ORIGEN

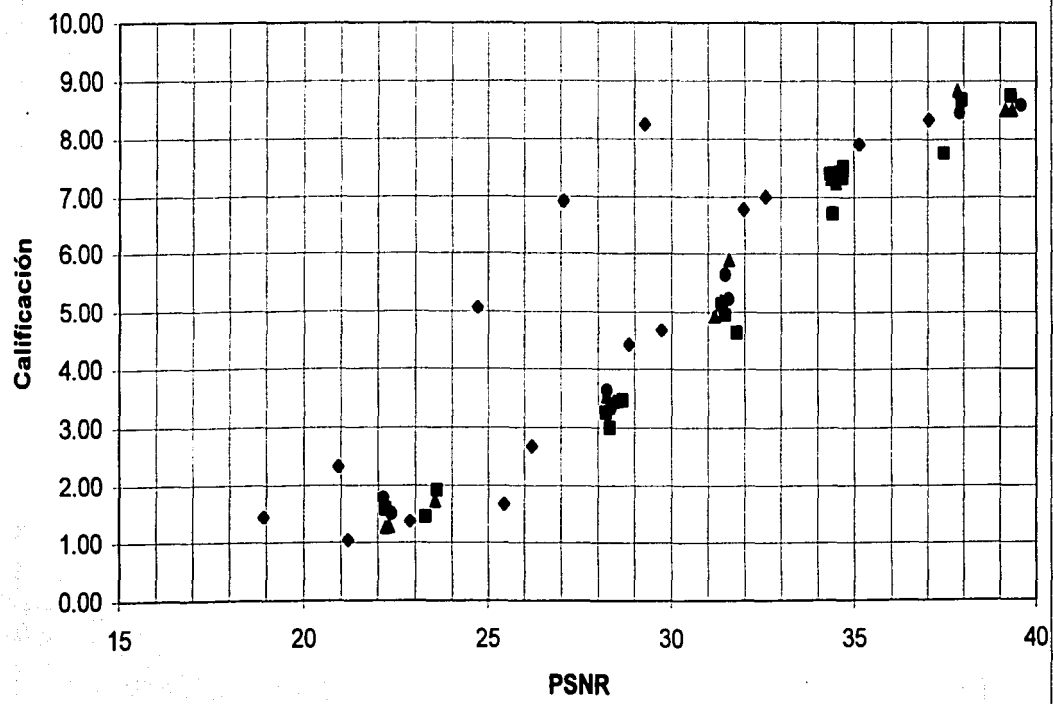


Grafica 5.7 Diagrama de dispersión de los resultados obtenidos

DIAGRAMA DE DISPERSION: CASA.RAW

Calidad objetiva vs Calidad subjetiva

- ◆ DB
- COIFLET
- ▲ CDF
- REV

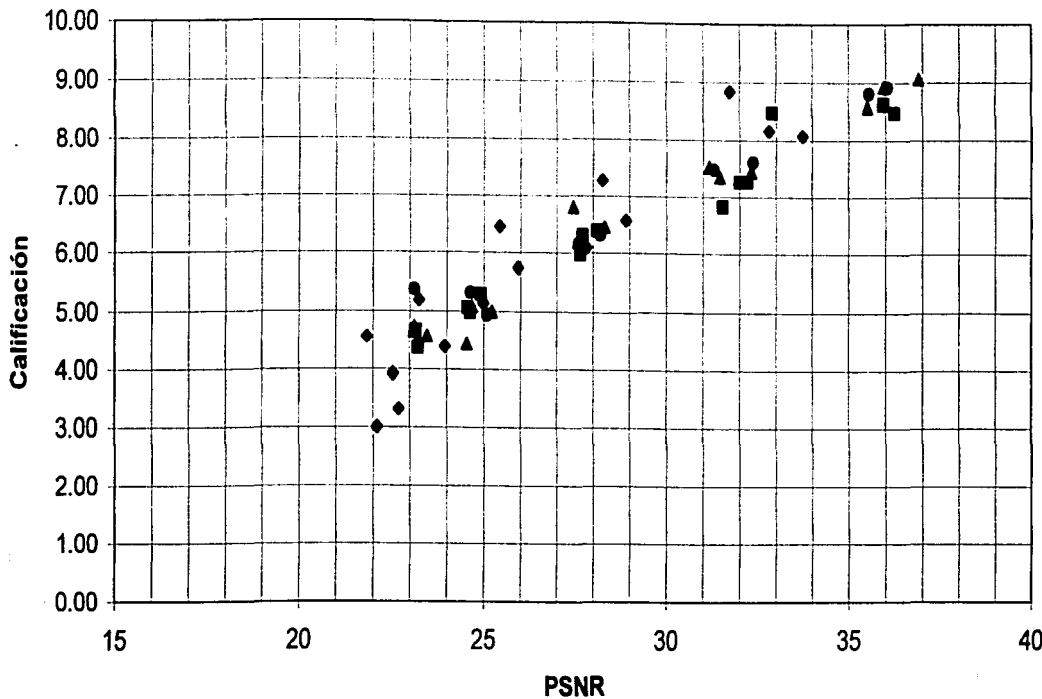


Grafica 5.8 Diagrama de dispersión de todos los sistemas wavelets para casa.raw

DIAGRAMA DE DISPERSION: BARBARA.RAW

Calidad objetiva vs Calidad subjetiva

- ◆ DB
- COIFLET
- ▲ CDF
- REV

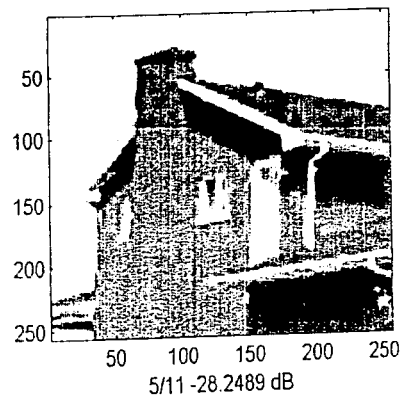
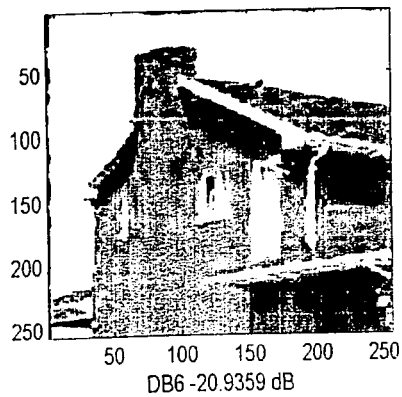
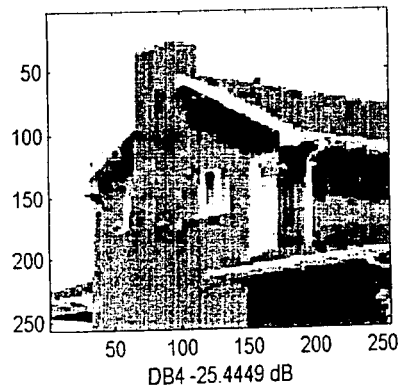
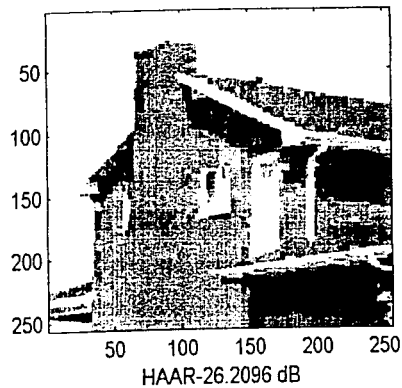


Grafica 5.9 Diagrama de dispersión de todos los sistemas wavelets para barbara.raw

RESERVA DE DERECHOS
TESIS CON
NOO SISILL

ESTA TESIS NO SALE
DE LA BIBLIOTECA

TASA: 0.125 bpp



TESIS CON
FALLA DE CUBIEN

Fig. 5.2(a) Imágenes comprimidas a una tasa de 0.125 bpp (1ª Parte)

TASA: 0.125 bpp

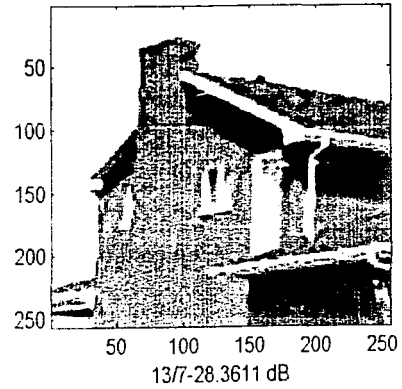
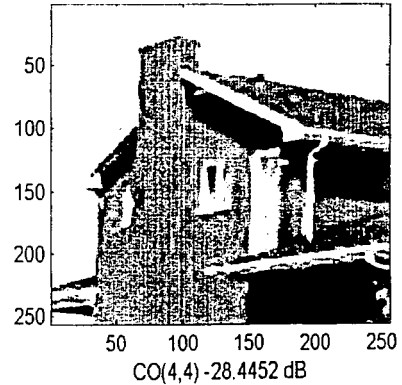


Fig. 5.2(b) Imágenes comprimidas a una tasa de 0.125 bpp (2ª Parte)

TESIS CON
FALLA DE ORIGEN

MEMORIA DE TESIS
CON
LISTA DE CONTENIDOS

TASA: 0.125 bpp

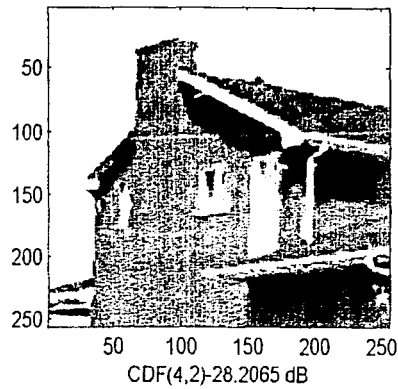
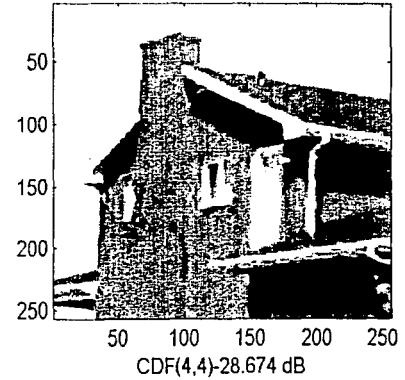
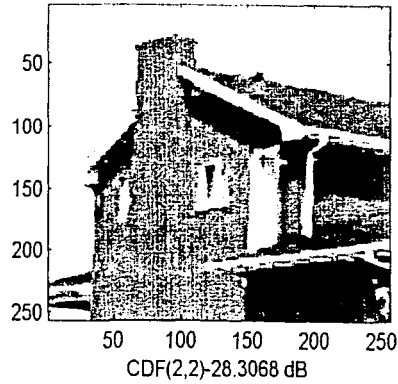


Fig. 5.2(c) Imágenes comprimidas a una tasa de 0.125 bpp (3ª Parte)

TASA: 0.125 bpp

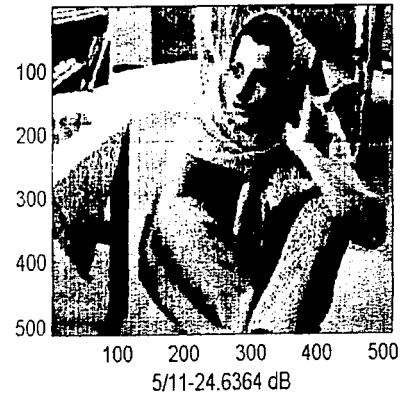
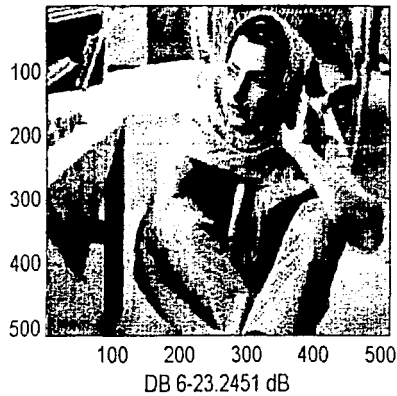
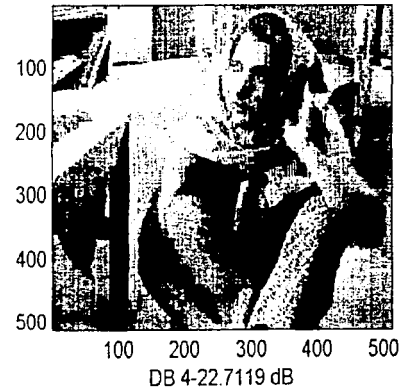
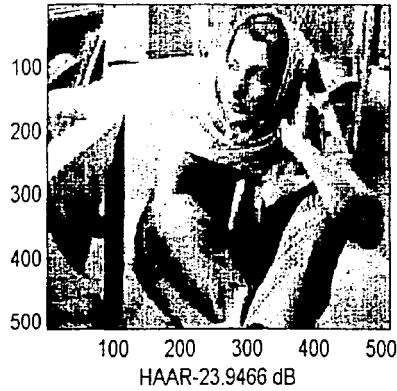
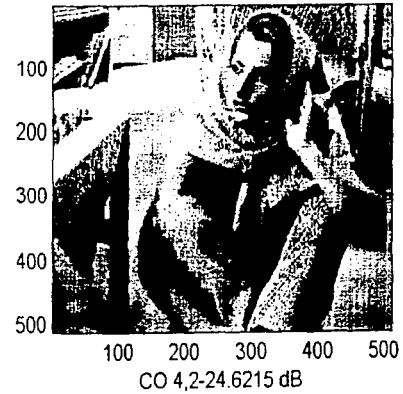
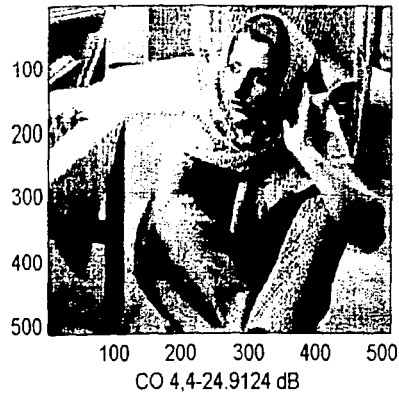


Fig. 5.3(a) Imágenes comprimidas a una tasa de 0.125 bpp (1ª Parte)

TESIS CON
FALLA DE ORIGEN

TASA:0.125 bpp



TESIS CON
SERIAL DE ORIGEN

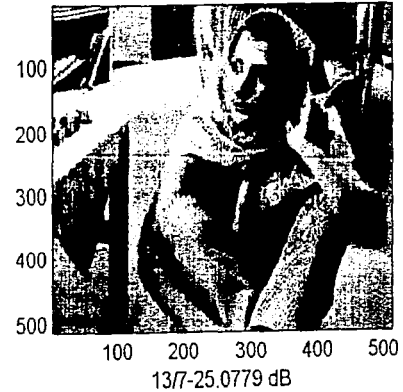
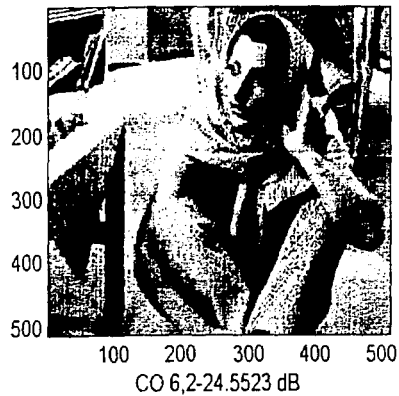
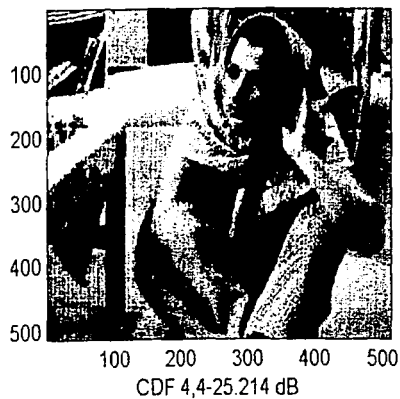
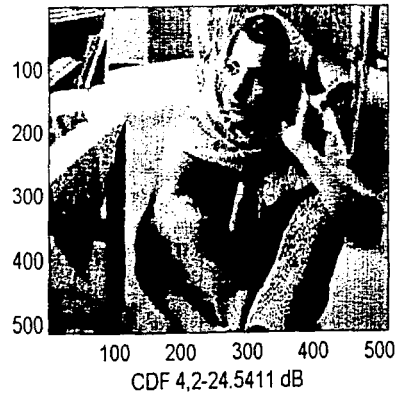
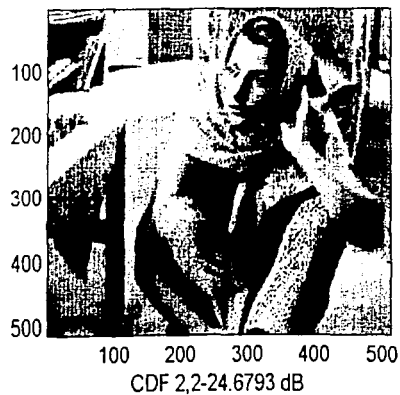


Fig. 5.3(b) Imágenes comprimidas a una tasa de 0.125 bpp (2ª Parte)

TASA: 0.125 bpp



TESIS CON
FALLA DE ORIGEN

Fig. 5.3(c) Imágenes comprimidas a una tasa de 0.125 bpp (3ª Parte)

Capítulo 6

CONCLUSIONES

Finalmente expondré las notas obtenidas con el análisis de los resultados experimentales; el propósito del capítulo es vincular los argumentos teóricos con lo obtenido en los experimentos, en cuanto a la teoría de wavelets como a los fundamentos del proceso de compresión.

6.1 Discusión de los resultados cuantitativos.

El rango de tasas de compresión seleccionado, nos centra en el uso de altas tasas de compresión, en las figuras 5.2 y 5.3 se muestran los resultados a una tasa de 0.125 bpp, sin embargo, para el caso de contar con recursos escasos, la tasa de 0.0625 bpp no es prohibitiva (dependiendo de la aplicación) representando una reducción considerable a la cantidad de información a ser manejada; en el caso de que los píxeles originalmente sean de 8 bits, esto corresponde aproximadamente a 99% menos de datos transmitidos o almacenados (la diferencia estriba en la cantidad de información de control que debe ser añadida). En la figura 6.1 (a) y (b) se presentan los mejores resultados cuantitativos, en este caso las mejores wavelets son la b-spline de (4,4) momentos nulos y la coiflet de (4,4), como se puede observar a pesar de sufrir una severa distorsión por la pérdida de detalles, la estructura de la casa se preserva y es posible distinguir el objeto que representa la imagen.

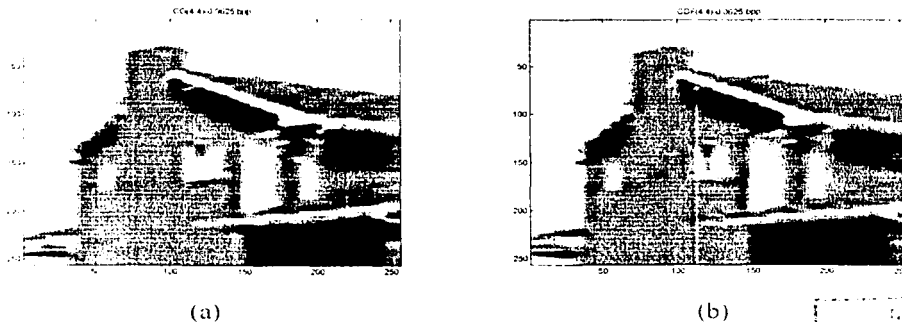
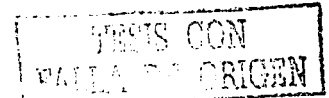


Fig 6.1 Casa a 0.0625 bpp.



A lo largo de la presentación de los resultados, se realizaron tablas separadas de cada imagen analizada, debido a que durante el desarrollo de los experimentos se encontraron diferencias entre el desempeño de las wavelets en cada caso, las cuales podrían pasar desapercibidas si se presentaran resultados globales. Esto se ve soportado, además, porque las características estadísticas de ambas imágenes son diferentes; mientras que la escena de la casa presenta algunas regiones homogéneas, en la imagen de Bárbara las texturas son predominantes. En la tabla 5.2 de los resultados

cuantitativos, se nota que en promedio los resultados obtenidos para la Casa son superiores a los de Bárbara, lo más interesante es que los sistemas se comportaron diferente, aunque en general podemos encontrar que:

- La B-spline, abreviada CDF(4,4), fue la de mejor desempeño a bajas tasas de compresión, mientras que la familia de Daubechies quedó en último lugar, siendo la Haar superior a las Daubechies de 4 y 6 momentos nulos. En el caso de las B-splines el mejor desempeño lo logra la wavelet donde el número de momentos nulos es igual en el análisis y en la síntesis, seguida por la coiflet que cumple con las mismas características.
- La wavelet Haar tiene dos propiedades relevantes, que solo en este caso conviven juntas en una wavelet real, es ortogonal y sus filtros de análisis y síntesis son simétricos, esto es un indicio de que estas propiedades deben ser más ponderadas que el número de momentos nulos del sistema. Otra punto es que la Haar tiene el soporte más corto entre las wavelets ortogonales pero no es apropiado a la aproximación de funciones suaves ya que tiene solamente un momento nulo.
- Si observamos las gráficas 5.1 y 5.2 que corresponden a la tabla ya mencionada las diferencias en el desempeño de las wavelets biortogonales se encuentran muy cerradas. Se puede observar notoriamente los grupos de wavelets biortogonales y ortogonales, siendo superiores los sistemas biortogonales, lo que podemos justificar por la simetría y, por lo tanto, la fase lineal de los filtros utilizados para implementarlos.
- Para la familia de coiflets, a bajas tasas de compresión, los mejores resultados se obtienen para el sistema donde el número de momentos nulos coinciden para el análisis y la síntesis, dejando en el último lugar de esta familia aquella con el mayor número de momentos nulos que corresponde a la coiflet (6.2).

Dentro del contexto general descrito anteriormente, hay que hacer los siguientes comentarios:

- En cuanto a la imagen de la Casa el peor desempeño lo tuvo invariablemente DB6, mientras que en el caso de Bárbara se presentó con mayor frecuencia que la de peor desempeño era DB4, esto indica que DB6 representa con mayor eficiencia los detalles de la imagen que DB4, aunque la diferencia no es grande.
- Cuando la tasa de compresión aumenta, para la Casa el mejor desempeño lo tiene la coiflet (4,4); en el caso de Bárbara lo tiene la coiflet (6,2), los sistemas reversibles también son buena opción, en el caso de la Casa es mejor la 5/11, la cual tiene un filtro más corto que la 13/7, esto no ocurre para Bárbara.

6.2 Discusión de los resultados cualitativos.

Los resultados cualitativos son menos determinantes que los cuantitativos, para identificar el conjunto de wavelets de mejor y peor desempeño, los datos se muestran dependientes de la tasa de compresión utilizada. Sin lugar a dudas las wavelets ortogonales son calificadas en este rubro como los sistemas más pobres, quedando en promedio un punto y medio debajo de los mejores resultados. Concentrándonos en la forma de onda; la familia Daubechies no se caracteriza por considerar esta propiedad en el diseño de la wavelet, incluso dada la naturaleza de la construcción de estos sistemas (basados en el número de los momentos nulos), las formas de onda son visiblemente irregulares.

Perceptualmente, la imagen de Bárbara motiva a concentrarse en los detalles que preserva la imagen después del proceso de compresión, aunque la estructura de ladrillo de la casa es también atractiva para evaluar la pérdida de nitidez, la Bárbara se muestra más cargada de información y más susceptible a perderla. En la figura 6.2 (a) y (b) se presentan los mejores resultados cualitativos a una tasa de 0.0625 bpp.

Ya se mencionaron algunos argumentos sobre la mayor ponderación que se le debe dar a la simetría de los filtros, en cuanto a la calidad subjetiva lograda por filtros simétricos el siguiente hecho refuerza la premisa anterior. Nuevamente las calificaciones más bajas la obtuvieron la familia Daubechies, pero no es la wavelet Haar la que prefirieron los usuarios en los experimentos sino la DB6, es decir la wavelet con el mayor número de momentos nulos, esto se puede explicar ya que como trabajamos a bajas tasas de bits un efecto marcado en las imágenes comprimidas con la Haar es que aparecen artefactos de bloque debido a la forma de onda de la wavelet, este fenómeno no se presenta en los otros sistemas ortogonales. Por lo que la wavelet Haar visualmente no es un opción atractiva específicamente porque solo tiene un momento nulo, sin embargo, la Haar no queda en último lugar, la DB4 tuvo más ocurrencias de ser calificada con el menor performance, y es cuando se enfatiza la relevancia de la forma de onda, pues como se puede observar en la figura 4.11, DB4 es la wavelet que presenta forma de onda menos suave. Incluso DB6 al tener mayor número de momentos nulos es más regular [C5]. También se puede notar en la tabla 5.3 que conforme se disminuye la tasa de compresión aumenta la distancia entre las calificaciones obtenidas por DB4 y DB6.

Para las tasas de compresión de 1 y 0.5 bpp las wavelets biortogonales de peor desempeño son la CDF (2,2) y la CDF (4,2), un resultado que a primera vista parecería factible es que entre las familias de Coiflets, B-Spline y Reversibles, la de mejor resultados fuera la familia spline ya que se diseñan con base en funciones cuyos coeficientes son binomiales, sin embargo, debido a las condiciones de ortogonalidad y por el orden de la función, la forma de onda de estas wavelets son modificadas e incluso para el caso de ordenes bajos (específicamente 2), la forma de onda en el análisis, y para el caso de la CDF (2,2) en la síntesis, es irregular – la forma de onda es triangular-. [Tabla 4.1]

En cuanto a los sistemas reversibles, el desempeño cualitativo mejoró incluso a las B-Splines, siendo una característica la forma de su onda, la cual se presenta en las figuras 4.9 y 4.10, como se puede observar en ellas, la función es más suave, aproximándose a las formas de onda gaussiana. En esta familia la de mejor desempeño es la 5/11, diferenciándose de la 13/7 por el tamaño de sus filtros, evidenciándose que con respecto a este factor un valor medio es el indicado

Otro aspecto que podemos notar en los resultados cualitativos son pequeñas discrepancias con lo obtenido con métricas cuantitativas. Iniciando con la wavelet Haar que en los resultado globales cuantitativamente superó a las demás wavelets ortogonales, en las calificaciones cualitativas no fue preferida por los efectos de bloque que presenta a bajas tasas de compresión consideradas. En el caso de Bárbara la wavelet CDF(4,4), que obtuvo una buena evaluación con las métricas objetivas, presenta la desventaja que pierde gran parte de los detalles de las texturas y la distorsión en los rasgos de la cara por lo que disminuyó su evaluación cualitativa; sin embargo, el desempeño que presenta permanece uniforme e incluso mejora comparativamente al disminuir la tasa de compresión.



Fig 6.2 Bárbara a 0.0625 bpp.

6.3 Conclusiones

El inminente uso del estándar JPEG 2000 y sus extensiones (como la Parte II que incluye la manipulación de la transformada wavelet) motiva el estudio de estas transformaciones y su impacto en el contexto del estándar, aunque se pueden encontrar más literatura sobre las wavelets reversibles, la decisión sobre su uso es completamente dependiente de la aplicación, ya que las wavelets irreversibles son exclusivas de la compresión con pérdida. Es importante recordar que la aproximación sucesiva que se logra mediante el codificador por entropía del JPEG 2000 nos permite, usando transformaciones reversibles, codificar una imagen con pérdida hasta lograr una compresión sin pérdida, si el proceso no es truncado prematuramente por falta de recursos.

Uno de los impactos del JPEG 2000 ha sido la introducción de las wavelets como herramienta de la codificación por transformada, abriendo la oportunidad no solo de explorar y explotar otros tipos de transformaciones, sino de mejorar los sistemas que actualmente conocemos. Por otro lado las wavelets como herramienta de análisis han tenido un desarrollo acelerado en las últimas dos décadas, hasta la reciente aparición de las wavelets de segunda generación.

Hasta el momento pocos han sido los estudios sobre las transformaciones wavelets en el contexto del estándar JPEG 2000, incluso ha sido mayor el interés sobre las transformaciones reversibles. El uso de diferentes wavelets en este contexto tuvo la finalidad de mejorar el performance de la compresión de las imágenes. Los resultados encontrados nos demuestran que el performance de la compresión con pérdida es altamente dependiente de la imagen analizada ya que ningún filtro se desempeñó uniformemente mejor que los otros en las imágenes procesadas. Esto hace pensar que la selección debería depender de la naturaleza estadística de la imagen a ser procesada y en esta tendencia han aparecido trabajos que defienden el uso de sistemas wavelets adaptivos. [D4] Sin embargo, es necesario recordar que en el nuevo estándar, la wavelet permanece constante en el proceso, por lo que sería necesario un pre-procesamiento que permitiera seleccionar la wavelet adecuada a la señal, lo que representaría un costo demastado alto para aplicaciones en tiempo real.

Aunque el performance global de la compresión depende de cada una de las etapas que lo conforman, la selección de la wavelet apropiada o del banco de filtros asociado tiene especial importancia, ya que aunque la cuantización sea adecuada y el codificador por entropía sea eficiente, si los filtros tienen un desempeño pobre, el codificador frecuentemente no proporcionará la ganancia suficiente para mantener la calidad de la imagen. La búsqueda de una wavelet adecuada para la compresión, se centró en el uso de funciones gaussianas o parientes cercanos, pero tomando el lineamiento del estándar, en el que se considera que para fines de eficiencia en la implementación el uso de filtros se limitaría a los FIR's, las funciones de soporte infinito no fueron incluidas en el estudio.

De la discusión de resultados podemos concluir que el conjunto de sistemas wavelets que presentaron mejor desempeño está formado por

- B-Spline CDF (4,4)
- Reversible 5/11
- Coiflet CO (4,4)

Estos sistemas tienen las siguientes propiedades en común:

- Son sistemas biortogonales
- Los soportes no son pequeños, alrededor de 10 puntos son tomados en el filtrado.

TESIS CON
 FALLA DE ORIGEN

- Tienen el mismo número de momentos nulos en el análisis y en la síntesis.
- El número de momentos nulos es el mayor de cada familia.
- En el caso de los resultados cualitativos la forma de onda es la más suave.

Finalmente los resultados mostrados nos dan herramientas para determinar los factores relevantes en el diseño de nuevas y más eficientes transformadas wavelets:

Ninguno de los sistemas wavelet presentó un rendimiento mejor en todos los casos, e incluso éste dependió de la imagen que se estaba procesando, es más, las variaciones en el PSNR entre las imágenes reconstruidas, dado que los parámetros del codificador permanecieron iguales, solo se puede atribuir a las características inherentes a la imagen original. De lo que se concluye que las características de la imagen y específicamente su regularidad es la principal variable de desempeño de la compresión, ya que cuando la resolución se incrementa, la tasa de decremento del error de aproximación $\|f - P_{n,l}f\|$ depende de la regularidad de f .

Una transformación ortogonal nos asegura que los coeficientes obtenidos no conservan redundancia, pero limita grandemente la forma de onda, haciéndola irregular. En el caso de los sistemas biortogonales esto se ve compensado al tener un sistema dual, una restricción que indican los resultados es que no es adecuado tener funciones wavelet de síntesis con una forma de onda regular deseada como es el caso de una función gaussiana, ya que esta selección afectó la regularidad y forma de onda de la wavelet de análisis primal. El sistema wavelet que muestra el mejor desempeño es aquella cuyo principal criterio de diseño es acercarse a un sistema ortogonal, en el cual los filtros de análisis y síntesis son iguales. Además de que para el caso de los sistemas CDF se cumpla con que los filtros de síntesis son funciones B-spline

Intentando establecer la prioridad de las características de las wavelets es preferible favorecer la simetría de los filtros, ya que es un factor que como se observó no sólo influye en la calidad subjetiva de la imagen sino en la objetiva también, esto es explicado por el retraso que en el dominio del espacio sufre la señal produciendo distorsiones (La distorsión que se presenta en los bordes es un desfase cuando el retardo de la señal de salida no es constante debido al uso de filtros de fase no lineal). Es mejor que el sistema sea implementado con filtros de fase lineal, lo que da preferencia a los sistemas biortogonales.

Entre más suave es la función wavelet de síntesis, es menor el número de efectos de bloque que se generan durante la compresión; la regularidad de la función ayuda a la visibilidad en la imagen porque los errores son de menor magnitud. En el caso de las wavelets de Haar y de la CDF (2,2) es evidente que, en las más altas tasas de compresión, imágenes con mejor calidad se obtienen con wavelets que son continuamente diferenciables que con wavelets discontinuas, razón por la cual es mejor aproximar las imágenes con señales suaves. En el diseño de wavelets algo de regularidad es deseada (el performance con no regularidad es pobre), aunque la regularidad más alta mejora la compresión, el performance no se ve afectado substancialmente.

Los momentos nulos en la etapa de análisis son necesarios porque gracias a estos la magnitud de los coeficientes wavelets disminuyen al aumentar la escala, logrando uno de los objetivos de la codificación por transformada: la compactación de la energía. Hablando desde el punto de vista de localización en la frecuencia, el desvanecimiento hacia frecuencias bajas corresponde al número de momentos nulos de la wavelet

En contraparte, la regularidad en la etapa de síntesis ayuda a aminorar los efectos de bloque, ya que permite la reconstrucción precisa de las señales en el sentido de que la expansión wavelet converge al espacio de regularidad al que pertenece la señal (Hölder o Sobolev).

Además estas dos propiedades tienen impacto desde el punto de vista numérico, ya que influyen directamente en la magnitud de los coeficientes de la transformación. Los momentos minimizados en la función de escala está relacionado a la eficiencia de la representación de los coeficientes de escala en altas resoluciones, ya que reduce la magnitud de estos coeficientes, además de que afecta en la simetría y concentración de las funciones de escala y wavelets.

La combinación de momentos minimizados de las funciones wavelet y escala, para la aproximación a las muestras de la señal original, dan mejores resultados que las wavelets que solo presentan momentos minimizados en su función wavelet; los ceros en los momentos de la función de escala no sólo mejora la aproximación de los coeficientes de la expansión, sino que provoca que la función de escala sea más simétrica. Por eso, la suavidad y la regularidad son propiedades que se deben buscar en la función de escala de la síntesis, convirtiéndose en parámetros claves en el diseño de los sistemas wavelets.

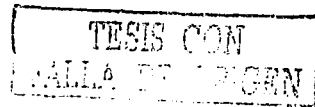
Del número de momentos nulos y el tamaño del soporte de las funciones base, siendo dos variables que compiten (ya que entre más regular es una función se requiere que los filtros asociados sean de mayor longitud), es preferible la segunda en el análisis y la primera en la síntesis. Esto se debe a que en el análisis las propiedades de localización (en el espacio principalmente) son más importantes mientras que en la síntesis el evitar la introducción de artefactos y el preservar los patrones perceptualmente relevantes en la imagen, permite proporcionar una buena calidad en la imagen reconstruida.

Al seleccionar una wavelet nos enfrentamos al compromiso existente entre el número de momentos nulos y el tamaño de soporte de la wavelet: si f tiene pocas singularidades aisladas y es muy regular entre singularidades, debemos seleccionar una wavelet con muchos momentos nulos para producir un gran número de coeficientes wavelets de magnitud pequeña; si la densidad de las singularidades se incrementa, sería mejor disminuir el tamaño de su soporte aún a costa de reducir el número de momentos nulos, ya que las wavelets cuando se traslapan en las singularidades crean coeficientes de amplitud grande. Con este hecho volvemos a converger en consideraciones de la función a representar para la selección de la wavelet adecuada.

En cuanto al tamaño del soporte ya se ha expuesto las razones al preferir los soportes pequeños, sin embargo dado el compromiso que se presenta con la forma de onda y el índice de regularidad los soportes de los filtros deben ser no menores a 5 coeficientes. Además no es posible aumentar el número de momentos nulos sin incrementar el tamaño del soporte de los filtros asociados.

La selectividad en frecuencia, directamente dependiente de las propiedades de localización de la wavelet, no es un factor determinante; sin embargo, debe considerarse que dado el principio de incertidumbre la selectividad en frecuencia conlleva a tener que ceder en la localización espacial. En este mismo sentido, entre mayor sea su métrica de regularidad en el dominio espacial, mejor será su selectividad en frecuencia.

Para el caso de las wavelets biortogonales, entre más parecidas sean las wavelets de análisis de las de la síntesis, el sistema tiende a ser un sistema ortogonal. Como hemos visto en los resultados entre más cercanos sean el número de momentos nulos en el análisis que en la síntesis, las formas de onda tienen semejante regularidad y se obtienen mejores resultados objetivos. Sin embargo en el



ámbito de los resultados cualitativos, el mejor desempeño lo obtienen los sistemas wavelets con mayor regularidad.

De las características analizadas, como ya se mencionó, la simetría es la más relevante, ya que el impacto en los bordes de la imagen es importante al eliminar por completo los efectos de bordes. En las características de los filtros esto se traduce en la característica de fase lineal, que sólo los filtros biortogonales nos pueden ofrecer.

Definitivamente el uso de sistema biortogonales es preferible a los ortogonales por la característica de simetría de los filtros, discutida a lo largo de este trabajo. La familia ortogonal de Daubechies es parametrizada por un entero que representa el número de momentos nulos de la función base y que es proporcional a la longitud del filtro implementado. Esta misma longitud está relacionada con el grado de suavidad y regularidad de la wavelet. En cuanto a las wavelets biortogonales un factor que mejora el desempeño en la compresión es que se seleccionen los tamaños de los filtros de análisis y de síntesis lo mas cercanos posible.

El caso más agradable sería que se encontraran familias wavelet biortonormales donde las wavelets primales (del análisis) y duales (de la síntesis) coincidieran y que las normas de la aproximación y de la función original fueran iguales. Además se ha visto que en el caso de las bases incondicionales aunque los coeficientes son únicos, las normas son solamente equivalentes mientras que en el caso de los frames, los coeficientes no son únicos pero pueden ser seleccionados de tal manera de que las normas sean iguales. Aunque para fines de compresión, los frames no pueden ser utilizados, el eliminar la restricción del soporte compacto nos permitiría considerar familias ortonormales, que si cumplen con buena localización espacio – frecuencia, y simplificar el diseño global del sistema, aunque con un costo en la implementación (por la introducción de filtrado recursivo).

Finalmente, en cuanto al procedimiento numérico que se empleó para la implementación de los sistemas wavelets, hay que notar que:

(i) El uso de algoritmos piramidales y wavelets permiten que aspectos del sistema visual humano sean inteligentemente tomados en cuenta para que la señal enmascare el ruido. La calidad perceptual de la imagen reconstruida es mejorada, aunque los cálculos de la compresión teórica no distinguen a este método en comparación con otros.

(ii) En el procesamiento de imágenes, el adoptar un esquema basado en “lifting” puede disminuir considerablemente la demanda computacional de la transformada. Además de esto el esquema proporciona otras ventajas entre las que se puede mencionar que permiten la transformación entero a entero y el cálculo inmediato de los coeficientes wavelet.

(iii) Aunque el tipo de cuantización permaneció invariante en los experimentos, la combinación de la transformación – cuantización sí afecta el desempeño global de la compresión y es una razón más para preferir que los filtros wavelets sean “suaves”. Cuando el filtro es irregular, su respuesta a escalón es oscilante, en el caso de los bordes, esta es la respuesta que esperamos y por lo tanto, es importante que sea lo más estable posible, de lo contrario, los coeficientes que obtengamos serán de gran magnitud, siendo más afectados por la cuantización y produciendo tasas de compresión menores.

TESIS CON
FALLA DE ORIGEN

6.4 Investigación futura

Existen varios lineamientos que se pueden seguir para continuar el estudio de los sistemas wavelets:

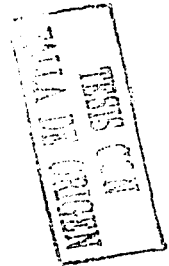
- Aunque se comprobó que las características de la imagen afectan el desempeño de la compresión, no se estableció con exactitud cuáles tienen mayor peso: el número de bits por píxel, la brillantez, alguno de sus componentes en frecuencia, su estadística, etc.
- Las metodologías para la construcción de sistemas wavelets han evolucionado desde que Meyer encontró sus wavelets, el diseño de la función wavelet, el análisis multi-resolución, hasta el esquema "lifting" permiten mayor flexibilidad en el diseño de las funciones base que cumplan con las condiciones impuestas, por lo que se abre el campo para la búsqueda de nuevos sistemas.
- El tercer punto se refiere a los filtros wavelets adaptivos, la búsqueda de algoritmos eficientes que permitan diseñar en línea los filtros adecuados para la señal procesada. La evidencia teórica y experimental permite suponer que la principal influencia sobre el desempeño de la compresión es debido a la regularidad de la señal a ser representada; se puede cambiar el enfoque del problema de procesamiento de señales buscando las funciones base apropiadas a nuestras señales, en lugar del enfoque tradicional en el que se trata de hacer que las señales sean apropiadas a las funciones base que utilizamos.

Aunque el presente trabajo intentó buscar la wavelet que se acopla mejor a las distintas clases de imágenes, una idea incipiente cambia radicalmente el enfoque de esta búsqueda; la posibilidad de diseñar dinámicamente la wavelet que se acople a las características de las señales que se están transmitiendo. Esta posibilidad se presenta gracias al desarrollo del esquema "lifting", el que nos permite diseñar diferentes wavelets incluso en el dominio del tiempo*, con lo que otra área de estudio que se abre es la adaptabilidad de los filtros "lifting" a las señales analizadas.

- El cuarto punto interesante es la propuesta de considerar los filtros IIR como opción para la implementación de transformaciones utilizando funciones con soporte infinito, pero con desvanecimiento rápido (como es el caso de las funciones gaussianas). Aunque el diseño de estos filtros resulta más complejo, si se logra obtener filtros IIR que aproximen aceptablemente las características atractivas de los filtros FIR (simetría y fase lineal), podremos considerar funciones con alto coeficiente de regularidad y que hasta el momento han permanecido marginadas y que no están acorde con la tendencia del diseño de wavelets de soporte compacto iniciada por Daubechies.
- La selección de wavelets y su relación con el proceso de cuantización puede ser mejorado si se toman en cuenta criterios perceptuales que eliminen la redundancia de los coeficientes o ayuden a ponderarlos de una manera provechosa para el observador. Antes de ordenar los coeficientes de los wavelet, éstos pueden ser ponderados por su visibilidad†, y más aún un conjunto de contextos y un modelo de probabilidad adecuado en la codificación por entropía podría ayudar a disminuir el número de bits por imagen.

* Hay que recordar que las primeras wavelets se diseñaron bajo el concepto de multi-resolución en el dominio de la frecuencia.

† Una medición psicofisiológica observable en la respuesta en frecuencia espacial del ojo humano.



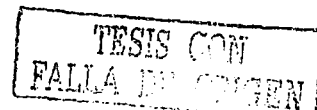
- ✓ Finalmente, en los recientes desarrollos surgen las multi-wavelets que remplazan una función de escala única de una variable real x con un vector de funciones de escala que satisfacen una versión matricial de la ecuación común de escala y , la función wavelet correspondiente se define de una manera matricial similar. Las ventajas que presenta este esquema es que las funciones wavelets son de tamaño de soporte mas corto, las funciones de escala que se involucran son simétricas lo que implica que las wavelets lo sean también, todas las traslaciones enteras de las funciones escala son ortogonales. Hay que recordar que es imposible tener funciones de escalas simétricas en el conjunto de las wavelets ortogonales, lo que hace necesario utilizar wavelets complejas o biortogonales para obtener la simetría. El soporte mas corto hace posible que los cálculos sean mejores

TESIS CON
FALLA DE ORIGEN

REFERENCIAS

- [A1] Akansu, Ali N et. al. *Wavelet and Subband Transforms: Fundamentals and Communication Applications*. IEEE Communications Magazine. December 1997. pp. 104-115.
- [A2] Antonini, Marc et. al. *Image Coding Using Wavelet Transform*. IEEE Transactions on Image Processing, Vol. 1, No. 2, Abril 1992. pp.205-220.
- [A3] Adams, M. & Kossentini, F. *Reversible Integer-to-Integer Wavelet Transforms for Image Compression*, IEEE Transactions on Image Processing, Vol. 9, No. 6, Junio 2000, pp. 1010-1024
- [A4] Adams, M. & Kossentini, F. *Low-complexity reversible integer-to-integer wavelet transforms for image coding*. Proc. IEEE Pacific Rim Conf. Communicatios, Computers, Signal Processing, Canada, Agosto 1999, pp 177-180.
- [B1] Bentley, P.M. & McDonnell, J.T.E. *Wavelet transforms: an introduction*. Electronics & Communication Engineering Journal. Agosto 1994. pp. 175-186.
- [B2] Burrus, C. Sidney et al. *Introduction to Wavelets and Wavelet Transforms: A Premier*. Prentice Hall. New Jersey 1998 pp.268.
- [C1] Wickerhouser, Mladen V. *Comparison of Picture Compression Methods: Wavelet, Wavelet Packet, and Local Cosine Transform Coding*. Wavelet Analysis and Its Applications, Vol.5, Edit. Academic Press, San Diego 1994, pp. 585-621.
- [C2] Calderbank, A.R. et. al. *Lossless Image Compression using integer to integer wavelet transforms*. Proceedings IEEE Int. Conf. Image Processing, Vol 1, Santa Barbara CA, Octubre 1997 pp. 596-599.
- [C3] Chen, Kuan-Fu et. al. *Analysis and Architecture Design of EBCOT for JPEG-2000*. Departamento de Ingeniería Eléctrica de la Universidad Nacional de Taiwán. Año 2001.
- [C4] Chui, Charles K. *An introduction to wavelets*. Edit Academic Press. Boston. 1992. 266 pp.
- [C5] Cohen, Albert & Ryan, Robert D. *Wavelets and Multiscale Signal Processing*. Serie: Applied Mathematics and Mathematical Computation. No: 11. Edit Chapman & Hall, 1995, 238 pp.
- [D1] Daubechies, Ingrid. *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol 61, Society for Industrial and Applied Mathematics. Philadelphia. 1992. 357 pp.
- [D2] Daubechies, Ingrid & Sweldens Wim, *Factoring Wavelet Transforms into Lifting Steps*, Program for Applied and Computational Mathematics, Princeton NJ, Nov. 1997.

- [D3] De Ridder, H. & Majoor, G.M.M. *Numerical category scaling: An efficient method for assessing digital image coding impairments*. Proceedings of the SPIE Conference on Human Vision and Electronic Imaging: Models, Methods and Applications, SPIE 1249, Santa Clara, California, 1990, pp. 65-77.
- [D4] DeVore, Ronald A. *Adaptive Wavelet Bases for Image Compression*. IEEE Transaction on Information Theory. Vol 38. 1995. pp. 197-219.
- [G1] Gopinath, R.A. & Burrus, C.S. *A Tutorial Overview of Filter Banks, Wavelet and Interrelations*, pp.104-107.
- [G2] Gray, Robert & Neuhoff, David. *Quantization* IEEE Transaction on Information Theory, Vol 44, No. 6, Octubre 1998. pp. 1-62.
- [G3] González, Rafael & Woods, Richard. *Digital Image Processing*, Addison Wesley, EEUU, 1992. 503 pp.
- [G4] Grangetto, Marco. et. al. *Finite Precision Wavelets for Image Coding: Lossy and Lossless Compression Performance Evaluation*. Departamento de electrónica del Politécnico de Torino. URL: helinet.polito.it/sasgroup. Año 2000.
- [J1] *JPEG 2000 Final Committee Draft Version 1.0*. ISO/IEC JTC 1/SC 29/WG. Abril 2000.
- [M1] Mallat, Stephane G. *A Wavelet Tour of Signal Processing*. Academic Press. San Diego. 1997. 312 pp.
- [M2] Marcellin, Michael & Fischer Thomas. *Trellis Coded Quantization of Memoryless and Gaussian-Markov Source*. IEEE Transactions on Communications, Vol. 38. No. 1 Enero 1990. pp. 82-93.
- [M3] Morgan, Alexander P., et. al., *A Gaussian Derivative Based Version of JPEG for Image Compression and Decompression*. IEEE Transactions on Image Processing, Vol. 7, No.9, Septiembre 1998, pp. 1311-1320.
- [M4] Meng, Hangying, et. al. *Performance of the Daubechies wavelet filters compared with other orthogonal transforms in random signal processing*. Proceedings of ICSP2000, pp. 333-336.
- [N1] Nestares, Oscar & Navarro, Rafael. *Efficient spatial-domain implementation of a multiscale image representation based on Gabor functions* Journal of Electronic Imaging, Vol 7, No. 1, Enero 1998, pp 166-173.
- [P1] Pennebaker W. et. al. *An overview of the basic principles of the Q-coder adaptive binary arithmetic coder*. IBM J. Res. Develop. Vol 36, Noviembre 1988, pp 717-726
- [P2] Parks, T.W., *Digital Filter Design*, Edit. John Wiley & Sons, New York. 1987. 342 pp.
- [R1] Reza, Ali M. *Wavelets Characteristics. What Wavelet Should I Use?*, Spire Lab. UWM, Octubre 1999.
- [R2] Ravela S. & Manmatha R. *Gaussian Filtered Representations of Images*. Department of Computer Science. University of Massachusetts. Technical Report.



- [R3] Roufs, J.A.J. & Boschman, M.C. *Visual comfort and performance*. Vision and Visual Dysfunction, Vol. 15, The Man-Machine Interface. Ed. Macmillan Press, Londres. 1991.
- [R4] Roufs, J.A.J. *Perceptual image quality: concept and measurement*. Philips Journal of Research 47, pp. 35-62.
- [S1] Sarkar, T.K., et. al. *A tutorial on Wavelets from an Electrical Engineering Perspective, Part I: Discrete Wavelet Techniques*. IEEE Antennas and Propagation Magazine, Vol. 40, No. 5, October 1998, pp. 49-70.
- [S2] Shapiro J. M. *Embedded Image Coding Using Zerotrees of Wavelet Coefficients*. IEEE Transactions on Signal Processing. Vol 41. Diciembre 1993 pp. 3445-3462
- [S3] Sweldens, Wim. *The Lifting Scheme: A new philosophy in biorthogonal wavelet constructions*. Proceedings SPIE Vol. 2569, pp. 68-79, Septiembre 1995.
- [S4] Sweldens, Wim. *Wavelets: What Next?*. AT&T Bell Laboratories. Bélgica. Enero 1996.
- [S5] Sweldens, Wim. *The Lifting Scheme: A custom-design construction of Biorthogonal Wavelets*. Technical Report, Industrial Mathematics Initiative, Department of Mathematics, University of South Carolina, Noviembre 1995.
- [S7] Said A. & Pearlman A. *A new fast and efficient image code on Set Partitioning in Hierarchical Trees*. IEEE Transactions on Circuits and Systems of Video Technology. Vol. 6. Junio 1996 pp. 243-250.
- [T1] Topiwala, Pankaj. *Wavelet Image and Video Compression*. The Kluwer International Series in Engineering and Computer Science. Boston 1998, 438 pp.
- [T2] Taubman, David. *High Performance Scalable Image Compression with EBCOT*. IEEE Transactions on Image Processing, Vol. 9, No. 7, Julio 2000, pp. 1158-1170.
- [U1] Unser Michael & Blu Thierry, *Why restrict ourselves to compactly supported basis functions?* Wavelets: Applications in Signal and Image Processing IX. Proceeding of SPIE Vol. 4478 (2001).
- [V1] Vetterli, Martin & Kovacevic, Jelena. *Wavelet and Subband coding*. Edit. Prentice Hall, New Jersey, 1995, 488 pp.

TESIS CON
FALLA DE ORIGEN

Resumen

El auge de las comunicaciones digitales, conduce a la necesidad de diseñar nuevos esquemas de compresión de imágenes que ahorren recursos tanto en el almacenamiento como en la transmisión de la señal, la cual debe estar disponible para el equipo que lo requiera. Para cumplir con el segundo requisito, el esquema debe acoplarse a un estándar que permita al equipo receptor de la señal, interpretarla correctamente.

El JPEG 2000 es el estándar emergente para imágenes estáticas de color y en niveles de grises; por un lado toma fundamentos del JBIG que se emplea para la codificación de imágenes en blanco y negro, y por otro está presente en el estándar MPEG 4 usado para la compresión de video. Sus principales características son:

- Sistema de compresión basado en transformada wavelet, en lugar de la DCT*.
- Codificación por planos de bits.
- Codificador por entropía basado en un árbol jerárquico que utiliza codificación aritmética de índices de contexto en lugar del tradicional código de Huffman.
- Maneja un esquema de transmisión progresiva que permite la interoperabilidad de los sistemas de comunicación.
- Soporte de escalabilidad de cantidad (SNR) y de escala.
- Codificación de ROIs (Regiones de interés) en una calidad mejor y menor distorsión que el resto de la imagen.
- Protección a la seguridad de las imágenes, permitiendo añadir información de derechos de autor y uso.

De este esquema la atención se centro en la selección apropiada del sistema wavelet (ortogonal o biortogonal) que permitiera alcanzar mejor desempeño tanto objetiva como subjetivamente.

El procedimiento para el análisis con wavelets es adoptar una función prototipo wavelet ψ , llamada wavelet madre o wavelet de análisis. El análisis temporal se realiza con una versión contraída y de alta frecuencia de la wavelet prototipo, mientras que el análisis en frecuencia se realiza con una versión dilatada y de baja frecuencia de la misma función. Esto equivale a analizar una señal a través de una ventana grande con la que podríamos notar las características generales, y con una ventana pequeña con la que lograríamos notar los detalles.

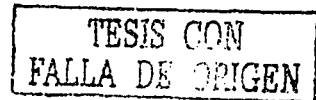
El tipo de la base ortonormal de $L^2(\mathbb{R})$ es:

$$\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k) \quad j, k \in \mathbb{Z}$$

con la expresión de descomposición:

$$f = \sum_{j,k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k}$$

* Transformada Coseno Discreta



El análisis multi-resolución introduce el concepto de wavelet padre ϕ , llamada también función de escala. Las traslaciones de esta función representan una base para el espacio que contiene la aproximación de la función f en la escala cero, llamada "aproximación en V_0 "; mientras que las traslaciones de ψ proporcionan el detalle que permite que la aproximación sea refinada a la siguiente escala, llamada "aproximación en V_1 ".

Las condiciones de ortogonalidad que se deben cumplir son:

$$\int \psi_{j,k}(x) \cdot \psi_{m,n}(x) dx = \begin{cases} 1 & j = m, k = n \\ 0 & \text{c.o.c.} \end{cases}$$

$$\int \phi_{j,k}(x) \cdot \phi_{m,n}(x) dx = \begin{cases} 1 & j = m, k = n \\ 0 & \text{c.o.c.} \end{cases}$$

$$\int \phi_{j,k}(x) \cdot \psi_{j,k}(x) dx = 0$$

Lo que se busca con la selección correcta de la wavelet es:

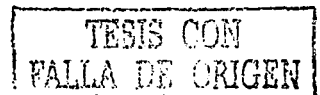
- Que la energía se concentre en el menor número posible de coeficientes, de tal manera que se pueda transmitir menor cantidad de información.
- Obtener coeficientes que sean adecuados para la cuantización[†].
- Encontrar una manera económica de enviar la información de los coeficientes.
- Simplicidad en el cálculo de la transformada.

Las transformadas reversibles o invertibles que transforman enteros a enteros son una pieza clave en la compresión sin pérdida definida por el estándar. Las imágenes de alta calidad como son el video con calidad de estudio, las imágenes médicas, los datos sísmicos, las imágenes de satélite y las imágenes digitalizadas de manuscritos con propósitos de preservación, son ejemplos de aplicaciones que demandan la compresión sin pérdida. El hecho de que en la transformación wavelet se obtengan coeficientes en punto flotante produce que la naturaleza de los datos enteros se pierda, para solventar este inconveniente la base para la creación de estas transformadas es un esquema relativamente nuevo llamado "lifting"; en el cual la transformación se realiza a través de actualizaciones y predicciones de la descomposición de la señal de entrada en la secuencia de muestras de índice par y las de índice impar. En el procesamiento de imágenes, el adoptar un esquema basado en lifting puede disminuir considerablemente la demanda computacional de la transformada, por lo que su uso no se limita a las transformaciones entero a entero.

Dado que no existen filtros FIR triviales, que sean ortonormales y de fase lineal con la propiedad de reconstrucción perfecta, considerando la regularidad (los únicos filtros simétricos de reconstrucción perfecta son los que corresponden a la base Haar); para preservar la fase lineal (lo que equivale a que los filtros FIR sean simétricos) es necesario relajar las condiciones de ortonormalidad y usar bases biortogonales.

La fórmula para la función de descomposición de las bases biortogonales es del tipo:

$$f(x) = \sum_{j \in \mathbb{Z}} 2^j \sum_{k \in \mathbb{Z}} \langle f, \psi_{j,k}(2^j x - k) \rangle \psi_{j,k}(2^j x - k)$$



[†] En este caso recordemos que el JPEG 2000 parte II considerará tanto la cuantización escalar como la cuantización TCQ.

Lo que implica que tenemos una wavelet de análisis φ y una wavelet de reconstrucción ψ y las condiciones de biortogonalidad quedan como:

$$\int \psi_{j,k}(x) \cdot \varphi_{m,n}(x) dx = \begin{cases} 1 & j = m, k = n \\ 0 & \text{c.o.c.} \end{cases}$$

$$\int \phi_{j,k}(x) \cdot \tilde{\varphi}_{m,n}(x) dx = \begin{cases} 1 & j = m, k = n \\ 0 & \text{c.o.c.} \end{cases}$$

$$\int \phi_{j,k}(x) \cdot \tilde{\psi}_{j,k}(x) dx = \int \tilde{\phi}_{j,k}(x) \cdot \psi_{j,k}(x) dx = 0$$

Ventajas

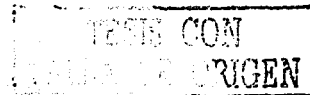
- Se obtienen más grados de libertad en el diseño de los filtros wavelet.
- Bancos de filtros que pueden estar constituidos de filtros de fase lineal
- Filtros/wavelets que pueden ser diseñados usando el esquema lifting.

La hipótesis que se sostuvo a lo largo de este estudio es que el seleccionar funciones wavelets relacionadas a las ondas gaussianas, puede ser un factor determinante para alcanzar un desempeño alto en la compresión. Para probar lo anterior, la metodología usada fue:

1. Selección de los factores que se considerarán relevantes en el desempeño de la transformada.
2. Selección de un conjunto de sistemas wavelets que de acuerdo a sus características aporten indicios sobre la importancia de la forma de onda de las funciones wavelet ψ y de escala ϕ .
3. Determinación de las variables en el estándar JPEG2000 para la inclusión de los sistemas wavelets seleccionados y realización de los experimentos correspondientes.
4. Evaluación objetiva y subjetiva de las imágenes obtenidas.
5. Conclusiones

Del punto 1, el conjunto de factores estudiados es:

- Ortogonalidad del sistema wavelet.
- Tamaño del soporte de las funciones wavelet.
- Número de momentos nulos² para la función wavelet.
- Regularidad de la función wavelet y de escala.
- Selectividad de frecuencia del sistema wavelet.
- Tamaño del filtros asociados al sistema wavelet.
- Fase de dichos filtros.
- Forma de onda de la wavelet.



Del punto 2, las familias estudiadas y sus características se presentan en el cuadro 4.1, mientras que en la tabla 5.1 se presenta la implementación lifting de cada una de ellas.

² En inglés "Vanishing Moments".

Las pruebas se llevaron a cabo con el Modelo de Verificación del JPEG 2000, versión 8 que tiene implementada parte de las funcionalidades de la parte II del estándar, principalmente la libre selección del kernel de la transformación. Para determinar completamente el experimento se fijaron las siguientes variables:

- Número de niveles de descomposición: La propiedad de aproximación sucesiva indica que conforme se añadan frecuencias más altas a la interpolación, se obtendrá una imagen de más alta resolución (esto equivale a generar más bandas o más señales de diferencia en la codificación sub-banda y codificación piramidal respectivamente) Sin embargo el tamaño de las imágenes que se procesarán no es lo suficientemente grande para seleccionar un número grande de niveles de descomposición. En este caso preferí utilizar el valor preestablecido en el estándar: 5 niveles.
- Estructura de descomposición: Se utilizó el esquema de Mallat.
- Tasas de compresión: Se utilizó la cuantización TCQ por lo que el usuario puede establecer a qué tasa se llevará a cabo la compresión, en este caso: 0.065, 0.125, 0.25, 0.50, 1 y 2 bpp.
- Imágenes sobre las que se trabajó: Son la Bárbara de 512 x 512 píxeles y la Casa de 256 x 256 píxeles (Figura 5.1). Ambas son imágenes en niveles de gris y tienen originalmente 8 bpp.
- Se evitó la división de la imagen en rejillas porque este proceso afecta directamente en los artefactos que presenta la imagen reconstruida, por lo que no se podrían discriminar de los artefactos que producen las transformadas, además dado el tamaño de las imágenes, no era primordial el uso de procesamiento en paralelo
- Los experimentos realizados utilizan el modo no reversible, el cual se explicó en el capítulo 3, sin embargo por condiciones del programa los filtros de análisis asociados a las wavelets se proporcionaron en su esquema lifting.

Para cumplir con el punto 4 de la metodología se utilizó la PSNR como métrica de la calidad objetiva de la imagen. Su definición es:

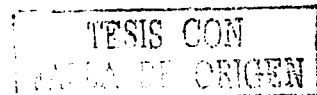
$$PSNR = 10 \log_{10} \left(\frac{2^p - 1}{MSE} \right) \quad (5.1)$$

donde

$$MSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (x[i, j] - \hat{x}[i, j])^2$$

Los experimentos que determinaron la calidad subjetiva se realizaron reuniendo a un grupo de 5 observadores no expertos; del conjunto de imágenes comprimidas con una tasa de bpp mayor a 2, se seleccionaron aleatoriamente 80 imágenes, al observador se le presentaron 100 imágenes con repeticiones; finalmente se les solicitó que asignaran a cada imagen una calificación entre 1 y 10.

Con las calificaciones obtenidas se obtuvieron promedios para cada tipo de wavelet y tasa de compresión. Adicionalmente se obtuvieron la calificación mínima y máxima para cada tasa de compresión, y el promedio por sistema wavelet para indicar el desempeño global del sistema.



Los resultados se muestran en la tabla 5.3. Las gráficas 5.3 y 5.4 muestran el desempeño subjetivo de las wavelets, mostrando la desviación estándar de cada promedio y finalmente en las gráficas 5.5 y 5.6 se presentan las comparaciones de estas métricas para la Casa y Bárbara, respectivamente.

Después del análisis de los resultados, podemos concluir que:

- ✓ El desempeño de la compresión depende principalmente de la regularidad de la señal, por lo que diseñar filtros acorde a las características de la imagen a procesar es una idea que puede desarrollarse en el futuro.
- ✓ La fase lineal de los filtros es imprescindible para evitar distorsiones en los bordes de la imagen, esta es la razón por la que los sistemas biortogonales son superiores a los ortogonales.
- ✓ Dentro de los sistemas biortogonales, los de mejor desempeño fueron aquellos en los que los filtros de análisis se asemejaban a los de síntesis, es decir, los sistemas que se aproximaban a los sistemas ortogonales.
- ✓ Por la naturaleza del diseño de los sistemas biortogonales, es claro que mientras más regular es la wavelet de análisis, más irregular es la wavelet de la síntesis. Sin embargo, la regularidad es una característica que se prefiere en la síntesis y para el caso de las calificaciones subjetivas la mejor calidad las obtuvieron la wavelet con la forma de onda más regular (los sistemas reversibles) y las wavelets que se basan en funciones b-spline; mientras que por calidad objetiva destacó la segunda. Por otro lado los momentos nulos, asociados con la propiedad de oscilación de la forma de onda es preferible que se presente en el análisis para asegurar la magnitud pequeña de los coeficientes wavelet.
- ✓ Las variables más influyentes en las calificaciones objetivas son: (i) Simetría de los filtros y (ii) Número de momentos nulos; mientras que para las calificaciones subjetivas son: (i) Simetría de los filtros, (ii) Regularidad de la función wavelet y (iii) Forma de onda de la wavelet.
- ✓ Aunque existe una correlación entre los resultados cualitativos y cuantitativos, existen claras diferencias como es el caso de la forma de onda, con la que mientras cualitativamente representa una ventaja la regularidad de la misma, mayor performance cuantitativo lo obtenemos cuando los sistemas biortogonales se aproximan a los ortogonales.
- ✓ Finalmente para el caso de las wavelets gaussianas, que no fueron consideradas en el estudio por no acoplarse a las condiciones del estándar, todavía resta el estudio de algoritmos rápidos para la transformación que permitan lograr una carga computacional apropiada a las aplicaciones de tiempo real, mientras que en el diseño resolver la problemática que surge al ser impuesto el uso de filtros IIR en lugar de filtros FIR.

TESIS CON
FALLA DE ORIGEN